

Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw mode!

Run on: January 11, 2003, 22:55:32 ; Search time 266 seconds
(without alignments)
194.722 Million cell updates/sec

Title: SEQ1-T-AT-12347_COPY_12336_12358

Perfect score: 23

Sequence: 1 999gaaatcttactttcgctcc 23

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 12

Total number of hits satisfying chosen parameters: 1343

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq

1: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1980.DAT:*

2: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1981.DAT:*

3: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1982.DAT:*

4: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1983.DAT:*

5: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1984.DAT:*

6: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1985.DAT:*

7: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1986.DAT:*

8: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1987.DAT:*

9: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1988.DAT:*

10: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1989.DAT:*

11: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1990.DAT:*

12: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1991.DAT:*

13: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1992.DAT:*

14: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1993.DAT:*

15: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1994.DAT:*

16: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1995.DAT:*

17: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1996.DAT:*

18: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1997.DAT:*

19: /S1DS2/gcdata/geneseq/geneseq-emb1/NA1998.DAT:*

20: /S1DS2/gcdata/geneseq/geneseq-emb1/NA2000.DAT:*

21: /S1DS2/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:*

22: /S1DS2/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:*

23: /S1DS2/gcdata/geneseq/geneseq-emb1/NA2002.DAT:*

24: /S1DS2/gcdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULTS

Result No.	Score	Query	Match Length	DB	ID	Description
1	16	69.6	700	22	AAH92847	Human inflammatory Human CDNA clone #
2	15	65.2	420	21	AAAT74522	Human CDNA #52 encoding Human immune/haema
3	15	65.2	462	23	AAS57386	Human mitochondria
4	15	65.2	588	22	AAK60492	Nucleotide sequence DNA encoding novel
5	15	65.2	630	22	AA070927	Human TBC-1 partia
6	15	65.2	946	22	AAF5428	Human nervous syst
7	15	65.2	1005	23	AAST7485	
8	15	63.2	99960	21	AAZ50905	
9	14	60.9	236	22	ABA122005	

ALIGNMENTS

Result No.	Score	Query	Match Length	DB	ID	Description
1	14	60.9	660	24	ABK76032	Bacillus lichenito oligonucleotide to
2	14	60.9	753	24	ABO47496	oligonucleotide fo
3	14	60.9	755	24	ABO47497	oligonucleotide fo
4	14	60.9	767	24	ABQ23634	oligonucleotide fo
5	14	60.9	767	24	ABQ23635	oligonucleotide fo
6	14	60.9	791	23	ABR5152	lRNA encoding novel
7	14	60.9	965	24	ABU87945	lRNA polymerase III
8	14	60.9	2622	22	ABU63737	RNA polymerase III
9	14	60.9	2697	22	AA160113	Rat sequence diffe
10	14	60.9	3610	23	ABU27970	Human polynucleot
11	14	60.9	4780	24	ABU91370	lrosophila melanog
12	14	60.9	6070	24	ABU21918	lrosophila melanog
13	14	60.9	6070	24	ABU9309	lrosophila melanog
14	14	60.9	6070	24	ABU32240	lrosophila melanog
15	14	60.9	6120	23	ABU29702	lrosophila melanog
16	14	60.9	6300	23	ABU69324	lrosophila melanog
17	14	60.9	6762	23	ABU04410	lrosophila melanog
18	14	60.9	17908	22	ABU37254	lrosophila melanog
19	14	60.9	17908	22	AAV28165	lRNA transcription
20	14	60.9	35099	19	AAV127112	Human immune syste
21	14	60.9	7353	24	AAS8958	Human chemically t
22	14	60.9	7353	24	AAS8363	Human chemically m
23	14	60.9	8047	24	AHN80325	lrosophila melanog
24	14	60.9	8047	24	ABD89245	lrosophila melanog
25	14	60.9	9762	23	ABU137254	lrosophila melanog
26	14	60.9	17908	22	AAV127112	lRNA transcription
27	14	60.9	17908	22	AAV28165	lRNA transcription
28	14	60.9	35099	19	AAV127112	Human immune syste
29	14	60.9	7353	24	AAS8958	Human chemically t
30	14	60.9	7353	24	AAS8363	Human chemically m
31	14	60.9	8047	24	AHN80325	lrosophila melanog
32	14	60.9	9762	23	ABD89245	lrosophila melanog
33	14	60.9	17908	22	AAV127112	lRNA transcription
34	14	60.9	17908	22	AAV28165	lRNA transcription
35	14	60.9	35099	19	AAV127112	Human immune syste
36	14	60.9	40267	24	AAS8958	Human chemically t
37	14	60.9	100301	24	ABO8176	Human osteoblast d
38	14	60.9	3011208	24	ABD69324	listeria innocua D
39	13	56.5	27	20	AAX77795	Staphylococcus sp.
40	13	56.5	97	16	AAT4854	Human gene signatu
41	13	56.5	100	16	AQ07384	Human type I stereo
42	13	56.5	163	16	AAT0244	Human gene signatu
43	13	56.5	197	21	AAC74424	Cat flea head and
44	13	56.5	209	21	AAC74424	Human secreted pro
45	13	56.5	211	20	AAX0460	Human secreted pro

RESULT 1 AAH92847/C

ID AAH92847 standard; DNA; 700 bp.

AC AAH92847;

DT 09-OCT-2001 (first entry)

DE Human inflammatory bowel disease related gene fragment lgr4172a.

XX Human inflammatory bowel disease; Crohn's disease; ulcerative colitis;

KW single nucleotide polymorphism; SNP; chromosome 19p13; paternity test;

chromosome 5p13-33; forensic test; gene therapy; ds.

XX Homo sapiens.

XX WO200142511-A2.

DE: Human immune/haematopoietic antigen encoding cDNA SEQ 1D NO:5552.
 XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ss.
 XX OS Homo sapiens.
 XX WO200157182-A2.
 PD 09-AUG-2001.
 XX
 PR 17-JAN-2001; 2001W0 US013541.
 PR 31-JAN-2000; 2000US-017965.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-018464.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218390.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225759.
 PR 14-AUG-2000; 2000US-0226368.
 PR 14-AUG-2000; 2000US-0226668.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227099.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229511.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0233400.
 PR 14-SEP-2000; 2000US-0233401.
 PR 14-SEP-2000; 2000US-0233663.
 PR 14-SEP-2000; 2000US-0233664.
 PR 14-SEP-2000; 2000US-0233965.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-023584.
 PR 27-SEP-2000; 2000US-023584.
 PR 27-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-023935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-024060.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0244674.
 PR 08-NOV-2000; 2000US-0245225.
 PR 08-NOV-2000; 2000US-0246375.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 08-NOV-2000; 2000US-0246627.
 PR 08-NOV-2000; 2000US-0246728.
 PR 08-NOV-2000; 2000US-0246732.
 PR 08-NOV-2000; 2000US-0246740.
 PR 08-NOV-2000; 2000US-0246747.
 PR 08-NOV-2000; 2000US-0246750.
 PR 08-NOV-2000; 2000US-0246757.
 PR 08-NOV-2000; 2000US-0246760.
 PR 08-NOV-2000; 2000US-0246767.
 PR 08-NOV-2000; 2000US-0246770.
 PR 08-NOV-2000; 2000US-0246773.
 PR 08-NOV-2000; 2000US-0246777.
 PR 08-NOV-2000; 2000US-0246778.
 PR 08-NOV-2000; 2000US-0246780.
 PR 08-NOV-2000; 2000US-0246781.
 PR 08-NOV-2000; 2000US-0246782.
 PR 08-NOV-2000; 2000US-0246783.
 PR 08-NOV-2000; 2000US-0246784.
 PR 08-NOV-2000; 2000US-0246785.
 PR 08-NOV-2000; 2000US-0246786.
 PR 08-NOV-2000; 2000US-0246787.
 PR 08-NOV-2000; 2000US-0246788.
 PR 08-NOV-2000; 2000US-0246789.
 PR 08-NOV-2000; 2000US-0246790.
 PR 08-NOV-2000; 2000US-0246791.
 PR 08-NOV-2000; 2000US-0246792.
 PR 08-NOV-2000; 2000US-0246793.
 PR 08-NOV-2000; 2000US-0246794.
 PR 08-NOV-2000; 2000US-0246795.
 PR 08-NOV-2000; 2000US-0246796.
 PR 08-NOV-2000; 2000US-0246797.
 PR 08-NOV-2000; 2000US-0246798.
 PR 08-NOV-2000; 2000US-0246799.
 PR 08-NOV-2000; 2000US-0246800.
 PR 08-NOV-2000; 2000US-0246801.
 PR 08-NOV-2000; 2000US-0246802.
 PR 08-NOV-2000; 2000US-0246803.
 PR 08-NOV-2000; 2000US-0246804.
 PR 08-NOV-2000; 2000US-0246805.
 PR 08-NOV-2000; 2000US-0246806.
 PR 08-NOV-2000; 2000US-0246807.
 PR 08-NOV-2000; 2000US-0246808.
 PR 08-NOV-2000; 2000US-0246809.
 PR 08-NOV-2000; 2000US-0246810.
 PR 08-NOV-2000; 2000US-0246811.
 PR 08-NOV-2000; 2000US-0246813.
 PR 08-NOV-2000; 2000US-0246827.
 PR 08-NOV-2000; 2000US-0246847.
 PR 08-NOV-2000; 2000US-0246852.
 PR 08-NOV-2000; 2000US-0246853.
 PR 08-NOV-2000; 2000US-0246854.
 PR 08-NOV-2000; 2000US-0246855.
 PR 08-NOV-2000; 2000US-0246856.
 PR 08-NOV-2000; 2000US-0246857.
 PR 08-NOV-2000; 2000US-0246858.
 PR 08-NOV-2000; 2000US-0246859.
 PR 08-NOV-2000; 2000US-0246860.
 PR 08-NOV-2000; 2000US-0246861.
 PR 08-NOV-2000; 2000US-0246862.
 PR 08-NOV-2000; 2000US-0246863.
 PR 08-NOV-2000; 2000US-0246864.
 PR 08-NOV-2000; 2000US-0246865.
 PR 08-NOV-2000; 2000US-0246866.
 PR 08-NOV-2000; 2000US-0246867.
 PR 08-NOV-2000; 2000US-0246868.
 PR 08-NOV-2000; 2000US-0246869.
 PR 08-NOV-2000; 2000US-0251989.

PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX PT Rosen CA, Barash SC, Ruben SM;
 XX FT DR WPI; 2001-483426/52.
 XX P-PSDB; AAMB7711.
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX PT
 PS Claim 1; SEQ ID NO 5552; 3071pp + Sequence listing; English.
 XX CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (1) amino acid sequences given in AAM82170 to AAM91921. (1) have cytostatic activity, and can be used in gene therapy and vaccine production. (1) proteins and Polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (1) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (1) by expressing inactive proteins or to supplement the patients own production of (1). Additionally, (1) Polynucleotides may be used to produce the secreted (1), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (1) proteins and Polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic derived cells. AAK64703 to AAK81694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM82169 represent sequences used in the exemplification of the present invention.
 XX SQ Sequence 588 BP; 183 A; 114 C; 128 G; 160 T; 3 other;
 Query Match 65.28%; Score 15; DB 22; length 588;
 Best Local Similarity 100.0%; Pred. No. 11; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GGC~~AA~~ACTCTACTTT 17
 Db 86 GGC~~AA~~ATCTTACTTT 72
 RESULT 5
 AAD07927/C
 ID AAD07927 standard; cDNA; 630 BP.
 XX AC AAD07927;
 XX DT 04-AUG-2001 (first entry)
 XX DE Human mitochondrial deformylase partial cDNA #2.
 XX KW Human; mitochondrial deformylase; neoplastic disease; cancer; AIDS; Acquired Immune Deficiency Syndrome; cell proliferation; cytostatic; nootropic; tranquiliser; antisense gene therapy; muscular dystrophy; amyotrophic lateral sclerosis; autoimmune disease; leukaemia; mood; nervous system disorder; anxiety; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; ss.
 XX OS Homo sapiens.
 XX
 FH Location/Qualifiers
 Key complement (1..630)
 CDS
 FT /*tag= a
 /product= "Human mitochondrial deformylase protein
 /fragment= #2"
 /transl_except= (pos:595..597, aa:Xaa)
 /transl_except= (pos:571..573, aa:Xaa)
 /transl_except= (pos:400..402, aa:Xaa)
 /transl_except= (pos:4..6, aa:Xaa)
 /transl_except= (pos:262..264, aa:Xaa)
 /transl_except= (pos:228..231, aa:Xaa)
 /transl_except= (pos:109..111, aa:Xaa)
 /transl_except= (pos:52..54, aa:Xaa)
 /transl_except= (pos:37..39, aa:Xaa)
 /transl_except= (pos:4..6, aa:Xaa)
 /note= "Xaa corresponds to in frame stop codon: GBS
 does not include start and stop codon" /partial
 XX PN WO200142431-A2.
 XX PD 14-JUN-2001.
 XX PT 01-DBSC-2000; 2000WO-EP12110.
 XX PR 08-DEC-1999; 99US-0169615.
 XX PR 11-OCT-2000; 2000US-0239106.
 XX PA (FARB) BAYER AG.
 XX P1 Ramakrishnan S;
 XX DR WPI; 2001-381664/40.
 XX DR P-PSDB; AAE03547.
 XX PT Treating neoplastic disease such as cancer by administering a reagent which modulates human mitochondrial deformylase activity or expression to modulate cell proliferation.
 XX PS Claim 10; Page 70; 73pp; English.
 XX CC The present sequence is a cDNA encoding human mitochondrial deformylase protein fragment. Mitochondrial deformylase is an enzyme which cleaves the formyl group from nascent formyl-methionine peptides. Mitochondrial deformylase is useful for treating neoplastic disease, AIDS, muscular dystrophy, amyotrophic lateral sclerosis, muscle wasting diseases and autoimmune diseases. Mitochondrial deformylase is useful for decreasing cell proliferation and is useful for treating cancers such as leukaemia, adenocarcinoma and other diseases involving increased levels of cell proliferation. Since mitochondrial deformylase is of importance to central and peripheral nervous system it is also useful as a target for treating nervous system disorders such as disorders of mood, anxiety disorders, disorders of thought and volition, disorders of sleep and wakefulness, neurodegenerative disorders such as Alzheimer's and Parkinson's disease. Mitochondrial deformylase DNA is also useful in antisense gene therapy.
 XX SQ Sequence 630 BP; 141 A; 176 C; 139 G; 171 T; 3 other;
 Query Match 65.28%; Score 15; DB 22; length 630;
 Best Local Similarity 100.0%; Pred. No. 11; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GGC~~AA~~ACTCTACTTT 16
 Db 338 GGC~~AA~~ATCTTACTTT 324
 RESULT 6
 AAF5428
 ID AAF5428
 XX AAF55428;
 XX AC AAF55428;
 XX DT 29-MAY-2001 (first entry)
 XX DE Nucleotide sequence of a human hydrolytic enzyme HYEN24.
 XX KW Human; hydrolytic enzyme; HYEN24; neurological disorder; cancer;
 KW immune system disorder; genetic disorder; cell proliferation disorder;
 KW epilepsy; ischemic cerebrovascular disease; stroke; Pick's disease;
 KW Huntington's disease; dementia; Parkinson's disease; multiple sclerosis;
 FT


```

FT
FT primer_bind
FT /note= "amplification of amplicon 99-20473"
FT /note= "complement (76771..76790)"
FT /tag= as
FT /bound_moiety= "Primer C11"
FT /note= "amplification of amplicon 99-20473"
FT 78292..78309
FT
FT primer_bind
FT /tag= at
FT /bound_moiety= "Primer B12"
FT /note= "amplification of amplicon 5-249"
FT /note= "complement (76704..78721)"
FT /tag= au
FT /bound_moiety= "Primer C12"
FT /note= "amplification of amplicon 5-249"
FT 81893..81912
FT
FT primer_bind
FT /tag= av
FT /bound_moiety= "Primer B13"
FT /note= "amplification of amplicon 99-20485"
FT /note= "complement (82353..82372)"
FT
FT primer_bind
FT /tag= aw
FT /bound_moiety= "Primer C13"
FT /note= "amplification of amplicon 99-20485"
FT 84392..84412
FT
FT primer_bind
FT /tag= ax
FT /bound_moiety= "Primer B14"
FT /note= "amplification of amplicon 99-20481"
FT /note= "complement (84909..84929)"
FT
FT primer_bind
FT /tag= ay
FT /bound_moiety= "Primer C14"
FT /note= "amplification of amplicon 99-20481"
FT 89745..89765
FT
FT primer_bind
FT /tag= az
FT /bound_moiety= "Primer B15"
FT /note= "amplification of amplicon 99-20480"
FT /note= "complement (90179..90198)"
FT
FT primer_bind
FT /tag= ba
FT /bound_moiety= "Primer C15"
FT /note= "amplification of amplicon 99-20480"
FT /note= "amplification of amplicon 99-20480"
FT 1431..1455
FT
FT misc_binding
FT /tag= bb
FT /bound_moiety= "Probe P2"
FT
FT
FT Query Match 65.2%; Score 15; DB 21; Length 99960;
FT Best Local Similarity 100.0%; Pred. No. 10; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT Oy 5 CAACTCTACTTTCG 19
FT Db 10209 CAAATCTTACATTCG 10195
FT
FT
FT RESULT 9
FT ABA12200/C
FT ID ABA12200 standard; cDNA; 236 bp.
FT XX
FT AC ABA12200;
FT XX
FT DT 23-JAN-2002 (first entry)
FT
FT Human nervous system related polynucleotide SEQ ID NO 1207.
FT
FT Human; nootropic; neuroprotective; cytostatic; dermatological; virucide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery; antiparkinsonian; antischlcking; antiaemic; antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective; antinflammatory; antiallergic; antidiabetic; antiluler; anticonvulsant; antifungal; antiparasitic; cardiot; immune disorder; anticonvulsant; antifungal; neurological disease; infection; nephrotropic; gene therapy; vaccine; ss. OS Homo sapiens.
FT XX
FT PN WO20159063-A2.
FT XX
FT PD 16-AUG-2001.

```

PR 29-SEP-2000; 2000US-0236369.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-02469208.
 PR 17-NOV-2000; 2000US-02469209.
 PR 17-NOV-2000; 2000US-02469210.
 PR 17-NOV-2000; 2000US-02469212.
 PR 17-NOV-2000; 2000US-02469213.
 PR 17-NOV-2000; 2000US-02469214.
 PR 17-NOV-2000; 2000US-02469215.
 PR 17-NOV-2000; 2000US-02469216.
 PR 17-NOV-2000; 2000US-02469218.
 PR 17-NOV-2000; 2000US-02469244.
 PR 17-NOV-2000; 2000US-02469245.
 PR 17-NOV-2000; 2000US-02469265.
 PR 17-NOV-2000; 2000US-02469266.
 PR 17-NOV-2000; 2000US-02469297.
 PR 17-NOV-2000; 2000US-02469299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-025671.
 PR 08-DEC-2000; 2000US-0254479.
 PR 08-DEC-2000; 2000US-0254856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Barash SC, Ruben SM:
 XX WPI: 2001-541565/60.
 DR P-PSB: ABB15874.

XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides, useful for preventing, diagnosing and/or treating nervous system cancers and metastases.
 PT
 XX
 PR Claim 1; SEQ ID NO. 1207; 1701pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABA11004-ABA21534) and proteins (ABBA4678-ABBA8001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (anti)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct-sequences.
 CC
 XX Sequence 236 BP; 69 A; 64 C; 39 G; 64 T; 0 other:
 Query Match 60.9%; Score 14; DB 22; Length 236;
 Best local Similarity 100.0%; Pred. No. 44; Mismatches 0;
 Matches 14; Conservative 0; Gaps 0;
 Oy 4 GCAAATCTGACTTT 17
 Db 43 GCAAATCTGACTTT 30

RESULT 10
 ARK76032
 ID ARK76032.standard; DNA; 660 bp.
 XX
 AC ARK76032;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Bacillus licheniformis genomic sequence tag (GST) #3423.
 XX
 KW differential gene expression; genomic sequencing tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds;
 XX US
 US
 XX
 XX
 PN WO200229113-A2.
 XX
 PD 11-APR-2002.
 XX
 PR 05-OCT-2001; 2001WO-US31437.
 XX
 PR 06-OCT-2000; 2000US-0680598.
 PR 27-MAR-2001; 2001US-27926P.
 XX
 PA (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 XX
 PI Berka R, Clausen IG;
 XX
 DR WPI; 2002-416684/44.
 XX
 PT Monitoring differential expression of several genes in first Bacillus cell relative to expression of same genes in one or more second Bacillus cells, by using substrate containing Bacillus genomic sequenced tag array -

xx
 PS Claim 4: SEQ ID NO 3323; 200pp; English.
 xx
 CC The invention describes a method of monitoring differential expression of
 CC genes in a first *Bacillus* cell relative to expression of the genes in
 CC other *Bacillus* cells, comprising hybridising labelled nucleic acid probes
 CC isolated from *Bacillus* cells to a substrate containing array of *Bacillus*
 CC genomic sequencing tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first *Bacillus* cell relative to expression of the same genes
 CC in one or more second *Bacillus* cells. The method is useful for monitoring
 CC global expression of several genes from a *Bacillus* cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which *Bacillus* cells adapt to changes in culture conditions.
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at
 CC <ftp://wipo.int/pb/published-pct-sequences>.
 xx
 SQ Sequence 660 BP; 224 A; 114 C; 137 G; 185 T; 0 other;
 xx
 Query Match 60.9%; Score 14; DB 24; Length 660;
 best local Similarity 100.0%; pred. No. 43;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 AACATCTTACTTCTCG 19
 Db 371 TACTCTTACTTCTCG 384
 xx
 RESULT 11
 ID ABQ47496
 XX ID ABQ47496 standard; DNA; 753 BP.
 XX ABQ47 96;
 AC ABQ47 96;
 DT 12-JUL-2002 (first entry)
 XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 34087.
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds;
 KW Homo sapiens;
 OS Homo sapiens;
 PN WO200218632-A2.
 XX PD 07-MAR-2002.
 XX PD 07-MAR-2002.
 XX PR 01-SEP-2001; 2001WO-EP10074.
 XX PR 01-SEP-2000; 2000DE-1043826.
 XX PR 05-SEP-2000; 2000DE-1044543.
 PA (EPIG-) EPIGENOMICS AG.
 XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX DR WPI: 2002-371829/40.
 XX PR Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 PT amplicons from chemically treated DNA -
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX

xx
 PS Claim 12; 56pp + sequence Listing; 56pp; German.
 xx
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis or side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc.; particularly by detecting mutations of similar nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC AB03410-AB054212 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 SQ Sequence 753 BP; 110 A; 77 C; 276 G; 290 T; 0 other;
 xx
 Query Match 60.9%; Score 14; DB 24; Length 753;
 best local Similarity 100.0%; pred. No. 43;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 TCTTACTTCTCGTC 22
 XX Db 56 TCTTACTTCTCGTC 43
 xx
 RESULT 12
 ID ABQ47497
 XX ID ABQ47497 standard; DNA; 753 BP.
 XX ABQ47 97;
 AC ABQ47 97;
 DT 12-JUL-2002 (first entry)
 XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 34088.
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds;
 KW Homo sapiens;
 OS Homo sapiens;
 PN WO200218632-A2.
 XX PD 07-MAR-2002.
 XX PF 01-SEP-2001; 2001WO-EP10074.
 XX PR 01-SEP-2000; 2000DE-1043826.
 XX PR 05-SEP-2000; 2000DE-1044543.
 PA (EPIG-) EPIGENOMICS AG.
 XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX DR WPI: 2002-371829/40.
 XX PR Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX

This invention describes a novel method for determining the degree of methylation of a particular cytosine in a motif 5'-CpG-3', present in a genomic sample of DNA. The sample is treated chemically to convert cytosine (C) but not methylated C, to uracil, then part of the genomic DNA that contains the target C is amplified to form a labeled amplicon. The amplicon is hybridised to two classes, each with at least one member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNPs); and (ii) for differentiation of cell or tissue types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. AB013410-AB04121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention.

Query	Match	Score	Length
Best	Local Similarity	14;	DB 24;
Matches	Conservative	100.0%	Pred. 43;
QY	ID	0;	Mismatches
Db	ABQ23634	0;	Indels
	TCTTA:TTCGGTC	0;	Gaps
RESULT	13		
ABQ23634	C		
ID	ABQ23634	standard;	DNA;
XX		767	BP.
AC	ABQ23634;		
XX			
DT	12-JUL-2002	(first entry)	
XX			
DE	Oligonucleotide for detecting cytosine methylation	SBQ	ID NO 10225.
XX			
KW	Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis; drug; side effect; cancer; central nervous system; cardiovascular; gastrointestinal; respiratory system; single nucleotide polymorphism; SNP; cell differentiation; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200218632-A2.		
XX			
PD	07-MAR-2002.		
XX			
PF	01-SEP-2001; 2001WO-EP10074.		
XX			
PR	01-SEP-2000; 2000DE-1043826.		
PR	05-SEP-2000; 2000DE-1044543.		
XX			
PA	(EPIG-) EPIGENOMICS AG.		
XX			
PI	Olek A, Pfepenbrock C, Berlin K, Guetig D;		
XX			
DR	WPI; 2002-371829/40.		
XX			
PT	Determining the degree of cytosine methylation in genomic DNA, useful for diagnosis and prognosis, comprises selective hybridization of amplicons from chemically treated DNA		
PT			
XX			
PS	Claim 12; 56pp + Sequence Listing; 56pp; German.		
XX			
CC	This invention describes a novel method for determining the degree of methylation of a particular cytosine in a motif 5'-CpG-3', present in a genomic sample of DNA. The sample is treated chemically to convert		

ID -
 XX ABQ23635 standard; DNA; 767 BP.
 XX AC
 XX DT ABQ23635;
 XX
 12-JUL-2002 (first entry)
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 10226.
 KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX OS Homo sapiens.
 XX PN WO200218632-A2.
 XX PD 07-MAR-2002.
 XX PF 01-SEP-2001; 2001WO-EP10074.
 XX PR 01-SEP-2000; 2000DE-1043826.
 XX PR 05-SEP-2000; 2000DE-1044543.
 XX PA (EPICG -) EPIGENOMICS AG.
 XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX DR WPI: 2002-371829/40.
 XX PT determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA.
 XX
 PS Claim 12; 56pp + sequence listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one

CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNP's); and (ii) for differentiation of cell or tissue types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. AbQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention.

XX Sequence 767 BP; 92 A; 318 C; 74 G; 283 T; 0 other;

Query Match 60.9%; Score 14; DB 24; Length 767;
Best local Similarity 100.0%; Pred. No. 43;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TCTTACTTTCGGTC 22
Db 8 TCTTACTTTCGGTC 21

RESULT 15

XX AAS86152 standard; cDNA; 791 BP.
XX AAS86152;
XX AAS86152;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #21956.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX KW Homo sapiens.
XX OS Homo sapiens.

PN WO200175067-A2.
XX P0 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.

PR XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.

PA XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.

DR P-PSDB; ABG21965.

XX New isolated Polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

XX Claim 1; SEQ ID No 21956; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS86152-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pdb/published_pct_sequences](http://wipo.int/pdb/published_pct_sequences).

XX Sequence 791 BP; 165 A; 202 C; 273 G; 151 T; 0 other;

Query Match 60.9%; Score 14; DB 23; Length 791;
Best local Similarity 100.0%; Pred. No. 43;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 CTTTACTTTCGGTC 23
Db 260 CTTTACTTTCGGTC 247

Search completed: January 11, 2003, 23:01:21
Job time : 314 secs

ON nucleic	- nucleic search, using sw model
Run on:	January 11, 2003, 23:02:06 ; Search time 2196 Seconds (without alignments) 169.625 Million cell updates/sec
Title:	SEQ1_T_AR_12347_COPY_12336_12358
Perfect score:	23
Sequence:	1 9999caaatcttacattcgctcc 23
Scoring table:	OLIGO-NUC
	Gapop 60.0 , Gapext 60.0
Searched:	16154066 seqs, 8097743376 residues
Word size	12
Total number of hits satisfying chosen parameters:	8301
Minimum DB seq length: 0	
Maximum DB seq length: 2000000000	
Post-processing:	Listing first 45 summaries
Database :	
	EST:*
1:	em_estba:*
2:	em_lesthium:*
3:	em_estin:*
4:	em_lestiu:*
5:	em_lestov:*
6:	em_lestpl:*
7:	em_estro:*
8:	em_htc:*
9:	qb_esti:*
10:	qb_est2:*
11:	qb_htc:*
12:	qb_est3:*
13:	qb_est4:*
14:	qb_est5:*
15:	em_lestiu:*
16:	em_lestom:*
17:	qb_gss:*
18:	em_gss_hum:*
19:	em_gss_inv:*
20:	em_gss_phn:*
21:	em_gss_vrt:*
22:	em_gss_fut:*
23:	em_gss_man:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_rod:*
Pred	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
	SUMMARIES
Result	Query
No.	Score
	Match Length
	DB
	ID
	Description
1	AG058557/C
2	AG058557
3	AG058557
4	AG058557
5	AG058557
6	AG058557
7	AG058557
8	AG058557
9	AG058557
10	AG058557
11	AG058557
12	AG058557
13	AG058557
14	AG058557
15	AG058557
16	AG058557
17	AG058557
18	AG058557
19	AG058557
20	AG058557
21	AG058557
22	AG058557
23	AG058557
24	AG058557
25	AG058557
26	AG058557
27	AG058557
28	AG058557
29	AG058557
30	AG058557
31	AG058557
32	AG058557
33	AG058557
34	AG058557
35	AG058557
36	AG058557
37	AG058557
38	AG058557
39	AG058557
40	AG058557
41	AG058557
42	AG058557
43	AG058557
44	AG058557
45	AG058557
46	AG058557
47	AG058557
48	AG058557
49	AG058557
50	AG058557
51	AG058557
52	AG058557
53	AG058557
54	AG058557
55	AG058557
56	AG058557
57	AG058557
58	AG058557
59	AG058557
60	AG058557
61	AG058557
62	AG058557
63	AG058557
64	AG058557
65	AG058557
66	AG058557
67	AG058557
68	AG058557
69	AG058557
70	AG058557
71	AG058557
72	AG058557
73	AG058557
74	AG058557
75	AG058557
76	AG058557
77	AG058557
78	AG058557
79	AG058557
80	AG058557
81	AG058557
82	AG058557
83	AG058557
84	AG058557
85	AG058557
86	AG058557
87	AG058557
88	AG058557
89	AG058557
90	AG058557
91	AG058557
92	AG058557
93	AG058557
94	AG058557
95	AG058557
96	AG058557
97	AG058557
98	AG058557
99	AG058557
100	AG058557
101	AG058557
102	AG058557
103	AG058557
104	AG058557
105	AG058557
106	AG058557
107	AG058557
108	AG058557
109	AG058557
110	AG058557
111	AG058557
112	AG058557
113	AG058557
114	AG058557
115	AG058557
116	AG058557
117	AG058557
118	AG058557
119	AG058557
120	AG058557
121	AG058557
122	AG058557
123	AG058557
124	AG058557
125	AG058557
126	AG058557
127	AG058557
128	AG058557
129	AG058557
130	AG058557
131	AG058557
132	AG058557
133	AG058557
134	AG058557
135	AG058557
136	AG058557
137	AG058557
138	AG058557
139	AG058557
140	AG058557
141	AG058557
142	AG058557
143	AG058557
144	AG058557
145	AG058557
146	AG058557
147	AG058557
148	AG058557
149	AG058557
150	AG058557
151	AG058557
152	AG058557
153	AG058557
154	AG058557
155	AG058557
156	AG058557
157	AG058557
158	AG058557
159	AG058557
160	AG058557
161	AG058557
162	AG058557
163	AG058557
164	AG058557
165	AG058557
166	AG058557
167	AG058557
168	AG058557
169	AG058557
170	AG058557
171	AG058557
172	AG058557
173	AG058557
174	AG058557
175	AG058557
176	AG058557
177	AG058557
178	AG058557
179	AG058557
180	AG058557
181	AG058557
182	AG058557
183	AG058557
184	AG058557
185	AG058557
186	AG058557
187	AG058557
188	AG058557
189	AG058557
190	AG058557
191	AG058557
192	AG058557
193	AG058557
194	AG058557
195	AG058557
196	AG058557
197	AG058557
198	AG058557
199	AG058557
200	AG058557
201	AG058557
202	AG058557
203	AG058557
204	AG058557
205	AG058557
206	AG058557
207	AG058557
208	AG058557
209	AG058557
210	AG058557
211	AG058557
212	AG058557
213	AG058557
214	AG058557
215	AG058557
216	AG058557
217	AG058557
218	AG058557
219	AG058557
220	AG058557
221	AG058557
222	AG058557
223	AG058557
224	AG058557
225	AG058557
226	AG058557
227	AG058557
228	AG058557
229	AG058557
230	AG058557
231	AG058557
232	AG058557
233	AG058557
234	AG058557
235	AG058557
236	AG058557
237	AG058557
238	AG058557
239	AG058557
240	AG058557
241	AG058557
242	AG058557
243	AG058557
244	AG058557
245	AG058557
246	AG058557
247	AG058557
248	AG058557
249	AG058557
250	AG058557
251	AG058557
252	AG058557
253	AG058557
254	AG058557
255	AG058557
256	AG058557
257	AG058557
258	AG058557
259	AG058557
260	AG058557
261	AG058557
262	AG058557
263	AG058557
264	AG058557
265	AG058557
266	AG058557
267	AG058557
268	AG058557
269	AG058557
270	AG058557
271	AG058557
272	AG058557
273	AG058557
274	AG058557
275	AG058557
276	AG058557
277	AG058557
278	AG058557
279	AG058557
280	AG058557
281	AG058557
282	AG058557
283	AG058557
284	AG058557
285	AG058557
286	AG058557
287	AG058557
288	AG058557
289	AG058557
290	AG058557
291	AG058557
292	AG058557
293	AG058557
294	AG058557
295	AG058557
296	AG058557
297	AG058557
298	AG058557
299	AG058557
300	AG058557
301	AG058557
302	AG058557
303	AG058557
304	AG058557
305	AG058557
306	AG058557
307	AG058557
308	AG058557
309	AG058557
310	AG058557
311	AG058557
312	AG058557
313	AG058557
314	AG058557
315	AG058557
316	AG058557
317	AG058557
318	AG058557
319	AG058557
320	AG058557
321	AG058557
322	AG058557
323	AG058557
324	AG058557
325	AG058557
326	AG058557
327	AG058557
328	AG058557
329	AG058557
330	AG058557
331	AG058557
332	AG058557
333	AG058557
334	AG058557
335	AG058557
336	AG058557
337	AG058557
338	AG058557
339	AG058557
340	AG058557
341	AG058557
342	AG058557
343	AG058557
344	AG058557
345	AG058557
346	AG058557
347	AG058557
348	AG058557
349	AG058557
350	AG058557
351	AG058557
352	AG058557
353	AG058557
354	AG058557
355	AG058557
356	AG058557
357	AG058557
358	AG058557
359	AG058557
360	AG058557
361	AG058557
362	AG058557
363	AG058557
364	AG058557
365	AG058557
366	AG058557
367	AG058557
368	AG058557
369	AG058557
370	AG058557
371	AG058557
372	AG058557
373	AG058557
374	AG058557
375	AG058557
376	AG058557
377	AG058557
378	AG058557
379	AG058557
380	AG058557
381	AG058557
382	AG058557
383	AG058557
384	AG058557
385	AG058557
386	AG058557
387	AG058557
388	AG058557
389	AG058557
390	AG058557
391	AG058557
392	AG058557
393	AG058557
394	AG058557
395	AG058557
396	AG058557
397	AG058557
398	AG058557
399	AG058557
400	AG058557
401	AG058557
402	AG058557
403	AG058557
404	AG058557
405	AG058557
406	AG058557
407	AG058557
408	AG058557
409	AG058557
410	AG058557
411	AG058557
412	AG058557
413	AG058557
414	AG058557
415	AG058557
416	AG058557
417	AG058557

PRIMERS

SEQUENCING: -21M13

LIBRARY

VECTOR : pK3145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

FEATURES source

1. -696

/organism="Pan troglodytes"

/db_xref="Taxon:9598"

/clone="PTB-045H18-F"

/clone="PTB-045H18-F"

/sex="Male"

/cell_type="Lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

BASE COUNT

164 a 206 c 168 g 156 t 2 others

ORIGIN

Query Match 78.3%; Score 18; DB 17; Length 696; Best Local Similarity 100.0%; Pred. No. 2.6; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCGAAATCTAGTTTCG 19

Db 387 GGCGAAATCTAGTTTCG 370

RESULT 2

AZ205659

LOCUS A2305659 667 bp DNA linear GSS 29-SEP-2000

DEFINITION IM0008K24F Mouse 10kb Plasmid UGGC1M library Mus musculus genomic

clone UGGC1M0008K24 F, DNA sequence.

ACCESSION A2305659

VERSION A2305659.1

KEYWORDS GSS

SOURCE house mouse

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 667)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah, Biomedical Polymers Research Bldg., 20 S. 2030 E., Salt Lake City, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddun@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Row: K Column: 24

Seq Primer: CGTGTAAACAGGAGGCCAGT

Class: Plasmid ends

High quality sequence stop: 667.

Location/Qualifiers

FEATURES source

1. -667

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="txon:10090"

/clone="UGGC1M0008K24"

/clone.lib="Mouse 10kb plasmid UGGC1M library"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42mv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/diaries/). The DNA was hydrodynamically sheared by repeated passage through a

RESULT 3

BF26081

LOCUS BF426081 391 bp mRNA 3'UTR cDNA clone GENEOME SYSTEMS CLONE ID: Gm-c1047-1024 5', mRNA sequence.

DEFINITION sr93f08.y1 Gm-c1047 Glycine max cDNA clone EST

ACCESSION BF426081

VERSION BF426081.1

KEYWORDS EST

SOURCE soybean

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine

REFERENCE 1 (bases 1 to 391)

AUTHORS Shiozaker,R., Keim,P., Vodkin,L., Erpfelding,J., Coryell,V., Khanna,Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rettner,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Watterton,R. and Wilson,R.

TITLE JOURNAL

COMMENT Unpublished (1993)

Contact: Shiozaker R/ public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com

Insert Length: 712 Std Error: 0.00

High quality sequence stop: 378.

Location/Qualifiers

FEATURES source

1. .391

/organism="Glycine max"

/db_xref="Taxon:3847"

/clone.lib="Gm-c1047"

/tissue-type="immature leaves (unfurled trifoliate) of greenhouse grown plants"

/clone.lib="Gm-c1047"

/note="Vector: pSFF01; Site_1: NotI; Site_2: SalI; This cDNA library was constructed from mRNA isolated from

immature leaves (unfurled trifoliate) of greenhouse grown plants that were 2 week old. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a NotI restriction site. Sait linkers adaptors were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPUR1 vector. The ligated cDNA fragments were transformed into E.coli ElectroporMax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign (email 1-vodkin@uiuc.edu).

BASE COUNT 125 a 61 c 80 g 125 t
ORIGIN Query Match 69.6%; Score 16; DB 12; Length 391;
Best Local Similarity 100.0%; Pred. No. 33; Mismatches 0;
Matches 16; Conservative 0; Indels 0; Gaps 0;
QY 1 GGGCAATCTACTT 16
Db 59 GGGCAATCTACTT 74

RESULT 4
AZ414906/c
LOCUS AZ414906 591 bp DNA linear GSS 03-OCT-2000
DEFINITION 1M0189F05R Mouse 10kb plasmid UGGCIM libr'ly Mus musculus genomic
clone UGGCIM0189F05 R, DNA sequence.
ACCESSION AZ414906
VERSION GSS.
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Mammalia: Eutheria: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciuromorpha: Muridae: Murinae: Mus.
REFERENCE 1 (Bases 1 to 591)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islan, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL Plasmid Inserts
COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dbunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0189 Row: F column: 05
Seq primer: CACACAGGAAACAGCTAAGACC
Class: plasmid ends
High quality sequence stop: 591.
FEATURES
source
1. .591
/organism="Mus musculus"
/strain="C5BL/6J"
/db_xref="taxon:10050"
/clone="UGGCIM0189F05"
/clone_libr="Mouse 10kb plasmid UGGCIM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD4Inv; Purified genomic DNA from M.
musculus C5BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource Laboratory (http://www.Jax.org/resources/documents/dmases/). The DNA was hydrodynamically sheared by repeated passage through a

RESULT 5
BJ141026/c
LOCUS BJ141026 785 bp mRNA linear EST 23-JAN-2002
DEFINITION Caenorhabditis elegans cDNA clone yk1165b07 3', mRNA sequence.
ACCESSION BJ141026
VERSION BJ141026.1
KEYWORDS EST.
TITLE
JOURNAL
COMMENT
ORGANISM Caenorhabditis elegans
Eukaryota: Metazoa: Chordata: Chromodorea: Rhabditida: Rhabditidae
; Rhabditidae: Pelecidae; Caenorhabditis.
1 (bases 1 to 785)
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.
A complementary view of the C.elegans genome
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute Of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
FEATURES
source
1. .785
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK1165b07"
/clone_libr="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/note="The AD-wrmcDNA library was generated with poly(A)+ RNA isolated from both hermaphrodite and male N2 worms of all larval stages, embryos, adults and dauer and the subsequent generation of cDNAs by poly(A) priming. The cDNAs were cloned into pPC86."

BASE COUNT 193 a 155 c 160 g 276 t 1 others
ORIGIN Query Match 69.6%; score 16; DB 13; Length 785;
Best Local Similarity 100.0%; Pred. No. 40; Mismatches 0;
Matches 16; Conservative 0; Indels 0; Gaps 0;

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 Polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|473211419|AF120721), a copy number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptorized mouse DNA was annealed to adaptored vector DNA, and transformed mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

/lab_host="DH10B (phage-resistant)"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Average insert size 2.1 kb."
 BASE COUNT 427 a 173 c 265 g 200 t
 ORIGIN

RESULT 9
 Query Match 69.6%; Score 16; DB 12; Length 1065;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 868 ATCTTACTTGGCTC 853

RESULT 10
 Query 4 GCAAATCTTACTTTC 18
 Definition 5', mRNA sequence.
 Locus AA233037
 Definition 2r46c09.r1 Soares, NHMPU_S1 Homo sapiens cDNA clone IMAGE:666443
 Accession AA233037
 Version AA233037.1 GI:1856030
 Keyword EST
 Source human.
 Organism Homo sapiens
 Taxonomy Homo sapiens
 Reference 1 (bases 1 to 275)
 Authors Hillier,L., Allen,M., Bowles,J., Dubuque,T., Geisel,G., Jost,S.,
 Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
 Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,
 Waterston,R., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Wilson,R. (1996)
 Title The WashU-HMM Mouse EST Project
 Journal Unpublished (1997)
 Comment Contact: Wilson KK
 WashU-HMM Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert length: 2067 Std Error: 0.00
 Seq primer: -28ml3 rev2 BT from Amersham
 High quality sequence stop: 258.
 FEATURES source
 1..275
 /organism="Homo sapiens"
 /db_xref="GDB:3428410"
 /clone="IMAGE:666448"
 /clone_id="Soares, NHMPU_S1"
 /tissue_type="pooled human melanocyte, fetal heart, and
 pregnant uterus"
 /lab_host="DH10B"
 /note="Organ: mixed (see below); Vector: pMT3D-Pac
 (Pharmacia) with a modified polylinker; Site_1: Not I;
 Site_2: Eco RI; Equal amounts of plasmid DNA from three
 normalized libraries (melanocyte 2NBHM, pregnant uterus
 NbHPU, and fetal heart NHH19W) were mixed, and ss circles
 were made in vitro. Following HAP purification, this DNA
 was used as tracer in a subtractive hybridization reaction.
 The driver was PCR-amplified cDNAs from pools of
 5,000 clones made from the same 3 libraries. The pools
 consisted of T.M.A.G.E. clones 26032-265223,
 340488-345479, and 484486-484479."
 BASE COUNT 92 a 44 c 56 g 83 t
 ORIGIN

RESULT 11
 Query Match 65.2%; Score 15; DB 9; Length 275;
 Best Local Similarity 100.0%; Pred. No. 1.2e-02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 107 GCAAATCTTACTT 93

BASE COUNT 65.2%; Score 15; DB 9; Length 101;
 ORIGIN

Query Match

LOCUS AA15557 276 bp mRNA linear EST 22-JAN-1998
 DEFINITION n53h10.r1 NCI_CGAP_Ew1 Homo sapiens cDNA clone IMAGE:1233571, mRNA
 sequence.

ACCESSION AA15557
 VERSION AA15557.1 GI:2727831
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 276) NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: crabps-r@mail.nih.gov
 Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
 Insert Length: 404 Std Error: 0.00
 Seq primer: -26n13 rev1 ET from Amersham.
 Location/Qualifiers

FEATURES source
 1. .276 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:233571"
 /clone.lib="NCI_CGAP_Ew1"
 /tissue_type="Ewing's sarcoma"
 /lab_host="DH10B"
 /note="vector: pAMP10; mRNA made from Ewing's sarcoma
 cDNA made by oligo-dT priming. Non-directionally cloned.
 Size selected on agarose gel, average insert size 600 bp.
 Reference: Krizman et al. (1996) Cancer Research
 56:530-533."

BASE COUNT 92 a 95 c 78 g 116 t
 ORIGIN 63 a 61 c 70 g 82 t

Query Match 65.2%; Score 15; DB 9; Length 276;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GGGCAAACTTACT 16
 Db 232 GGGCAAATCTTACTT 218

RESULT 12
 A1823498/c
 DEFINITION A1823498
 ACCESSION A1823498
 VERSION A1823498.1 GI:5444169
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 384) NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: crabps-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 Insert Length: 1062 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 378.
 Location/Qualifiers

FEATURES source
 1. .384 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2481039"
 /clone.lib="NCI_CGAP_Lu19"

BASE COUNT 92 a 95 c 78 g 116 t
 ORIGIN 63 a 61 c 70 g 82 t

Query Match 65.2%; Score 15; DB 9; Length 381;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GGGCAAACTTACTT 16
 Db 340 GGGCAAATCTTACTT 326

RESULT 13
 A1991677/c
 DEFINITION A1991677
 ACCESSION A1991677
 VERSION A1991677.1 GI:5838669
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 384) NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: crabps-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 Insert Length: 1062 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 378.
 Location/Qualifiers

FEATURES source
 1. .384 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2481039"
 /clone.lib="NCI_CGAP_Lu19"

CDNA Library preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution by: Washington University Genome Sequencing Center
 found through the T.M.A.G.E. Consortium/LN1 at:
 www.bio.lnl.gov/bbpr/image/image.html
 Insert Length: 437 Std Error: 0.00
 Seq primer: -40UP from Gibco.
 Location/Qualifiers

FEATURES source
 1. .381 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone.lib="NCI_CGAP_Kid11"
 /lab_host="DH10B"
 /note="Organ: kidney; vector: pTT73D-Pac (Pharmacia) with
 a modified polylinker; site_1: Not 1; Site_2: Eco RI;
 Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
 prepared, and ss circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (cloneIDs 1322776-1233911, 1456007-1456775, and
 150052-102855). Subtraction by Bento Soares and M.
 Fatima Bonaldo."
 Location/Qualifiers

CDNA Library preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution by: Washington University Genome Sequencing Center
 found through the T.M.A.G.E. Consortium/LN1 at:
 www.bio.lnl.gov/bbpr/image/image.html
 Insert Length: 437 Std Error: 0.00
 Seq primer: -40UP from Gibco.
 Location/Qualifiers

GenCore version 5.1.3
copyright (c) 1993 - 2003 Compugen Ltd.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

OM nucleic - nucleic search, using sw model
Run on: January 11, 2003, 23:01:32 ; search time 2901 Seconds
(without alignments)
230.736 Million cell updates/sec

Title: SFG1-T-AT-12347_COPY_12336_12358
Perfect score: 23
Sequence: 1 999gcaatcttacttgcgtcc 23
Scoring table: OLIGO.NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 12

Total number of hits satisfying chosen parameters: 15901

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Listing first 45 summaries

Database : GenBank:*

1: qb_ba: *
2: qb_htg: *
3: qb_in: *
4: qb_on: *
5: qb_ov: *
6: qb_sts: *
7: qb_ph: *
8: qb_p1: *
9: qb_pr: *
10: qb_ro: *
11: qb_sy: *
12: qb_un: *
13: qb_y1: *
14: em_ba: *
15: em_fun: *
16: em_hum: *
18: em_in: *
19: em_mu: *
20: em_on: *
21: em_or: *
22: em_ov: *
23: em_Pat: *
24: em_ph: *
25: em_p1: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_v1: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_p1: *
35: em_htg_r1: *
36: em_htg_main: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

ALIGNMENTS

RESULT 1
AC074203
LOCUS AC074203
DEFINITION Homo sapiens chromosome 22q11 clone cos4, complete sequence.
ACCESSION AC074203
VERSION AC074203.3 GI:9625348
KEYWORDS HTG
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 42479)
AUTHORS Yang, L., Hu, P., Wu, H., Tapia-Paez, I. and Roe, B. A.
TITLE Homo sapiens Chromosome 22q11 Cosmid cos4
JOURNAL Unpublished
Pred. No. is the number of results predicted by chance to have a

----- Project Information
 Center project name: HumDraft1
 Center clone name: CRC-22/C10
 ----- Summary Statistics
 Sequencing vector: PCR products; 100% of reads
 Assembly program: phrap; version 0.990329
 Consensus quality: 10728 bases at least Q40
 Consensus quality: 110501 bases at least Q30
 Consensus quality: 111810 bases at least Q20
 Insert size: 113327; sum-of-contigs
 Quality coverage: 6.70x in Q20 bases; sum-of-contigs

 NOTE: This is a 'working draft' sequence. It currently consists of 18 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 1 15681 contig of 15681 bp in length
 15782 30275 contig of 14494 bp in length
 30376 40528 contig of 10153 bp in length
 40629 50205 contig of 9577 bp in length
 50306 59617 contig of 9312 bp in length
 59718 66750 contig of 7033 bp in length
 66851 73847 contig of 6997 bp in length
 73948 79644 contig of 5697 bp in length
 79745 85733 contig of 5488 bp in length
 85754 85653 contig of 5909 bp in length
 91342 91241 contig of 5078 bp in length
 96520 100644 contig of 4125 bp in length
 100745 102923 contig of 2179 bp in length
 103024 105901 contig of 2927 bp in length
 106051 108938 contig of 2888 bp in length
 109039 111577 contig of 2539 bp in length
 111678 113487 contig of 1810 bp in length
 113588 115027 contig of 1440 bp in length
 Sequence updated (26-May-2000).
 * NOTE: This is a 'working draft' sequence. It currently consists of 18 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 * 1 15681: contig of 15681 bp in length
 * 15682 30275: contig of 14494 bp in length
 * 10276 30375: gap of 100 bp
 * 40629 40528: contig of 10153 bp in length
 * 40529 40628: gap of 100 bp
 * 50205 50205: contig of 9577 bp in length
 * 50206 50305: gap of 100 bp
 * 50306 59617: contig of 9312 bp in length
 * 59618 59717: gap of 100 bp
 * 59718 66750: contig of 7033 bp in length
 * 66751 66850: gap of 100 bp
 * 66851 73847: contig of 6997 bp in length
 * 73848 73947: gap of 100 bp
 * 73948 79644: contig of 5697 bp in length
 * 79645 79744: gap of 100 bp
 * 79745 85653: contig of 5909 bp in length
 * 85654 85733: gap of 100 bp
 * 85754 91241: contig of 5488 bp in length
 * 91242 91341: gap of 100 bp
 * 91342 96419: contig of 5078 bp in length
 * 96420 96519: gap of 100 bp
 * 96520 100644: contig of 4125 bp in length
 * 100645 100744: gap of 100 bp
 * 100745 102923: contig of 2179 bp in length
 * 102924 103023: gap of 100 bp

Sequence updated (26-May-2000).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 29 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 16627: contig of 16627 bp in length
 * 16628 16727: gap of 100 bp
 * 16728 30620: contig of 13893 bp in length
 * 30521 30720: gap of 100 bp
 * 30721 42439: contig of 11719 bp in length
 * 42440 42549: gap of 100 bp
 * 42540 53421: contig of 10882 bp in length
 * 53422 53521: gap of 100 bp
 * 53522 64748: contig of 11227 bp in length
 * 64749 64848: gap of 100 bp
 * 64849 76919: contig of 12071 bp in length
 * 76920 77019: gap of 100 bp
 * 77020 85523: contig of 8504 bp in length
 * 85524 85623: gap of 100 bp
 * 85624 94708: contig of 9085 bp in length
 * 94709 94808: gap of 100 bp
 * 94809 100076: contig of 11719 bp in length
 * 104077 104176: gap of 100 bp
 * 104177 112629: contig of 8453 bp in length
 * 112630 112729: gap of 100 bp
 * 112730 119309: contig of 6580 bp in length
 * 11931 119409: gap of 100 bp
 * 119410 126696: contig of 7287 bp in length
 * 126697 126796: gap of 100 bp
 * 126797 132422: contig of 5626 bp in length
 * 132423 132524: gap of 100 bp
 * 132523 1327920: contig of 5398 bp in length
 * 138021 138020: gap of 100 bp
 * 138021 142190: contig of 4170 bp in length
 * 142191 142290: gap of 100 bp
 * 142291 142583: contig of 2993 bp in length
 * 142584 145383: gap of 100 bp
 * 145384 145550: contig of 3167 bp in length
 * 148551 148651: gap of 100 bp
 * 148651 151542: contig of 2775 bp in length
 * 151426 151525: gap of 100 bp
 * 151526 151543: contig of 3907 bp in length
 * 151533 155532: gap of 100 bp
 * 158018 158018: contig of 2486 bp in length
 * 158119 160661: contig of 2591 bp in length
 * 160662 160761: gap of 100 bp
 * 160762 163242: contig of 2481 bp in length
 * 163243 163342: gap of 100 bp
 * 163343 165254: contig of 1912 bp in length
 * 165255 165334: gap of 100 bp
 * 165335 167624: contig of 2270 bp in length
 * 167625 167724: gap of 100 bp
 * 169399 169498: contig of 1674 bp in length
 * 169499 171146: contig of 1648 bp in length
 * 171147 171246: gap of 100 bp
 * 171247 173044: contig of 1798 bp in length
 * 173045 173144: gap of 100 bp
 * 173145 174332: contig of 1188 bp in length
 * 174333 174432: gap of 100 bp
 * 174433 175582: contig of 1150 bp in length.

FEATURES source

1 .175582
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q
 /clone="CTC-270C21"

FEATURES source

1 .175582
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q
 /clone="CTC-270C21"

RESULT 4

AC109505	AC109505	63669 bp	DNA	Linear	HTG 04-FEB-2002
LOCUS	DEFINITION	Mus musculus clone RP23-155M20, LOW-PASS SEQUENCE SAMPLING.			
ACCESSION	VERSION	AC109505	AC109505.1	GI:18483516	
KEYWORDS					
SOURCE					
ORGANISM					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
REFERENCE					
1 (bases 1 to 63669)					
AUTHORS					
Birren, B., Linton, L., Nusbaum, C. and Lander, E.					
TITLE					
Mus musculus, clone RP23-155M20					
JOURNAL					
Unpublished					
REFERENCE					
2 (bases 1 to 63669)					

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgatier, B., Brown, A., Camarata, J., Campolino, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Colpins, S., Colymore, A., Cook, A., Cooke, P., Dearalland, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Gailan, J., Gaidyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Huime, I., Iiev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamacaress, R., Landers, T., Lehoczyk, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Menaud, L., Minova, T., Mlenga, V., Murphy, T., Naylor, J., Nyuyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Rettig, R., Rieback, M., Riley, R., Risc, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vasiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A., and Zody, M.

TITLE JOURNAL
COMMENT
RESEARCH
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
Center project name: L20871
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIKR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L20871
Center clone name: 155_M_20

 * NOTE: This record contains 80 individual sequencing reads that have not been assembled into contigs. Runs of N are used to separate the reads and the order in which they appear is completely arbitrary. Low-pass sequence sampling is useful for identifying clones that may be gene-rich and allows overlap relationships among clones to be deduced. However, it should not be assumed that this clone will be sequenced to completion. In the event that the record is updated, the accession number will be preserved.

1 700: contig of 700 bp in length
 * 701 800: gap of 100 bp
 * 801 1500: contig of 700 bp in length
 * 1501 1600: gap of 100 bp
 * 1601 2254: contig of 654 bp in length
 * 2255 2354: gap of 100 bp
 * 2355 3058: contig of 704 bp in length
 * 3059 3158: gap of 100 bp
 * 3159 3869: contig of 711 bp in length
 * 3870 3969: gap of 100 bp
 * 3970 7067: contig of 683 bp in length
 * 4684 4783: gap of 100 bp
 * 4784 5482: contig of 699 bp in length
 * 5483 5582: gap of 100 bp
 * 5583 6284: contig of 702 bp in length
 * 6285 6384: gap of 100 bp
 * 6385 7067: contig of 683 bp in length
 * 7068 7167: gap of 100 bp
 * 7168 7867: contig of 700 bp in length
 * 7868 7967: gap of 100 bp
 * 7968 8651: contig of 684 bp in length
 * 8652 8751: gap of 100 bp
 * 8752 9470: contig of 719 bp in length
 * 9471 9570: gap of 100 bp
 * 10264: contig of 694 bp in length

10265 10364: gap of 100 bp
 * 10365 11045: contig of 681 bp in length
 * 11046 11145: gap of 100 bp
 * 11146 11864: contig of 700 bp in length
 * 11865 11964: gap of 100 bp
 * 11965 12666: contig of 702 bp in length
 * 12667 12765: gap of 100 bp
 * 12767 13471: contig of 705 bp in length
 * 13472 13571: gap of 100 bp
 * 13572 14250: contig of 679 bp in length
 * 14251 14350: gap of 100 bp
 * 14351 15046: contig of 696 bp in length
 * 15047 15146: gap of 100 bp
 * 15147 15828: contig of 682 bp in length
 * 15829 15928: gap of 100 bp
 * 15929 16629: contig of 701 bp in length
 * 16630 16729: gap of 100 bp
 * 16730 17424: contig of 695 bp in length
 * 17425 17524: gap of 100 bp
 * 17525 18216: contig of 692 bp in length
 * 18217 18316: gap of 100 bp
 * 18317 19010: contig of 694 bp in length
 * 19011 19110: gap of 100 bp
 * 19111 19821: contig of 711 bp in length
 * 19822 19921: gap of 100 bp
 * 19922 20627: contig of 706 bp in length
 * 20628 20728: gap of 100 bp
 * 20728 21433: contig of 706 bp in length
 * 21434 21533: gap of 100 bp
 * 21534 22238: contig of 705 bp in length
 * 22239 22339: gap of 100 bp
 * 22339 22038: contig of 700 bp in length
 * 23039 23138: gap of 100 bp
 * 23139 23837: contig of 699 bp in length
 * 23938 23938: gap of 100 bp
 * 24633 24632: contig of 695 bp in length
 * 24733 24732: gap of 100 bp
 * 25427 25526: gap of 100 bp
 * 25527 26223: contig of 697 bp in length
 * 26224 26323: gap of 100 bp
 * 26324 27022: contig of 699 bp in length
 * 27023 27122: gap of 100 bp
 * 27123 27825: contig of 703 bp in length
 * 27826 27925: gap of 100 bp
 * 27926 28632: contig of 707 bp in length
 * 28633 28732: gap of 100 bp
 * 28733 28428: contig of 696 bp in length
 * 289429 29528: gap of 100 bp
 * 29529 30222: contig of 694 bp in length
 * 30223 30323: gap of 100 bp
 * 30323 31013: contig of 691 bp in length
 * 31014 31113: gap of 100 bp
 * 31114 31806: contig of 693 bp in length
 * 31807 31908: gap of 100 bp
 * 31907 32605: contig of 699 bp in length
 * 32606 32705: gap of 100 bp
 * 32706 33034: contig of 689 bp in length
 * 33035 33491: gap of 100 bp
 * 33495 33187: contig of 693 bp in length
 * 33495 34188: gap of 100 bp
 * 34188 34281: gap of 100 bp
 * 34288 34989: contig of 702 bp in length
 * 34990 35059: gap of 100 bp
 * 35090 35798: contig of 709 bp in length
 * 35799 35898: gap of 100 bp
 * 35899 36597: contig of 699 bp in length
 * 36598 36691: gap of 100 bp
 * 36698 37398: contig of 701 bp in length
 * 37399 37498: gap of 100 bp
 * 37499 38207: contig of 709 bp in length
 * 38208 38307: gap of 100 bp
 * 38308 39006: contig of 699 bp in length
 * 39007 39105: gap of 100 bp

12767 13471: contig of 705 bp in length
 * 13572 13572: gap of 100 bp in length
 * 14251 14350: contig of 100 bp
 * 15047 15146: gap of 100 bp
 * 15147 15828: contig of 682 bp in length
 * 15829 15928: gap of 100 bp
 * 15929 16629: contig of 701 bp in length
 * 16630 16729: gap of 100 bp
 * 16730 17424: contig of 695 bp in length
 * 17423 17524: gap of 100 bp
 * 17525 18216: contig of 692 bp in length
 * 18217 18316: gap of 100 bp
 * 18317 19010: contig of 694 bp in length
 * 19011 19110: gap of 100 bp
 * 19111 19821: contig of 711 bp in length
 * 19822 19921: gap of 100 bp
 * 19922 20627: contig of 706 bp in length
 * 20628 20727: gap of 100 bp
 * 20728 21433: contig of 706 bp in length
 * 21434 21533: gap of 100 bp
 * 21534 22238: contig of 705 bp in length
 * 22239 22338: gap of 100 bp
 * 22339 23038: contig of 700 bp in length
 * 23039 23138: gap of 100 bp
 * 23139 23837: contig of 699 bp in length
 * 23838 23937: gap of 100 bp
 * 23938 24632: contig of 695 bp in length
 * 24633 24732: gap of 100 bp
 * 24733 25426: contig of 694 bp in length
 * 25427 25526: gap of 100 bp
 * 25527 26223: contig of 697 bp in length
 * 26224 26323: gap of 100 bp
 * 26324 27122: contig of 699 bp in length
 * 27123 27122: gap of 100 bp
 * 27822 27922: contig of 703 bp in length
 * 27926 28632: contig of 707 bp in length
 * 28633 28732: gap of 100 bp
 * 28733 29428: contig of 696 bp in length
 * 29429 29529: gap of 100 bp
 * 29529 30222: contig of 694 bp in length
 * 30223 30322: gap of 100 bp
 * 30323 31013: contig of 691 bp in length
 * 31114 31113: gap of 100 bp
 * 31807 31806: contig of 693 bp in length
 * 31907 32065: contig of 699 bp in length
 * 32606 32705: gap of 100 bp
 * 32706 33394: contig of 689 bp in length
 * 33395 33494: gap of 100 bp
 * 33495 34187: contig of 693 bp in length
 * 34188 34287: gap of 100 bp
 * 34288 34989: contig of 702 bp in length
 * 34990 35089: gap of 100 bp
 * 35090 35798: contig of 709 bp in length
 * 35799 35898: gap of 100 bp
 * 35899 36597: contig of 699 bp in length
 * 36598 36697: gap of 100 bp
 * 36698 37398: contig of 701 bp in length
 * 37399 37499: gap of 100 bp
 * 37499 38207: contig of 709 bp in length
 * 38208 38307: gap of 100 bp
 * 38308 39006: contig of 699 bp in length
 * 39007 39106: gap of 100 bp
 * 39107 39803: contig of 697 bp in length
 * 39804 39903: gap of 100 bp
 * 39904 40623: contig of 720 bp in length
 * 40624 40723: gap of 100 bp
 * 40724 41415: contig of 692 bp in length
 * 41416 41515: gap of 100 bp
 * 41516 42212: contig of 697 bp in length

Query Match Best Local Similarity 73.9%; Score 17; DB 2; Length 63669;
 Matches 17; Conservative 0; Mismatches 0; Insets 0; Gaps 0;

Qy 2 GGGCAATTCTTACITTC 18
 Db 1341 GGGCAATTCTTACITTC 1325

RESULT 6
 AC124530
 LOCUS AC124530 177182 bp DNA linear HTG 21-JUL-2002
 DEFINITION Mus musculus chromosome 11Nk clone Rb23-286D7, WORKING DRAFT
 ACCESSION AC124530
 VERSION AC124530.2 G1:21413484
 KEYWORDS HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Murida; Murinae; Mus;
 REFERENCE 1 (bases 1 to 177182)
 AUTHORS McPherson, J. D. and Waterston, R. H.
 TITLE The sequence of *Mus musculus* clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 177182)
 AUTHORS McPherson, J. D. and Waterston, R. H.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 REFERENCE 3 (bases 1 to 177182)
 AUTHORS McPherson, J. D. and Waterston, R. H.
 TITLE Direct Submission
 JOURNAL Submitted (21-JUL-2002) Genome Sequencing Center, 4444 Forest Park
 COMMENT On Jul 21, 2002 this sequence version replaced g1:21426654.
 * ----- Genome Center -----
 Center: Washington University Genome Sequencing Center

RESULT	7	2	GGCAAACTCTACTTTC	18
LOCUS	AC130947/c			
DEFINITION	Rattus norvegicus clone CH230-199D23, *** SEQUENCING IN PROGRESS			
ACCESSION	AC130947			
VERSION	AC130947.1			
KEYWORDS	HTG; HTGS_PHASE1;			
SOURCE	Norway rat.			
ORGANISM	Rattus norvegicus			
MATERIAL	Karyotype; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
REFERENCE	1 (bases 1 to 20953)			
AUTHORS	Muzny, D., Marie, M., Metzker, M., Lee, M., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alshiroos, S., Amin, A., Anguiano, D., Anyalebechi, V., Avioli, A., Ayodele, J., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, O., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buahy, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavares, J., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockell, R., Cox, M., Coyne, M., Creel, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Andrade, C., Dederich, D., Delgado, O., Denson, S., Derramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Hayes, K., Egan, A., Escott, M., Eugene, C., Evans, C. A., Faits, T., Fan, G., Fernandez, S., Finley, M., Flanq, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Gutierrez, W., Guerava, W., Guaratine, P., Haal, M., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylik, P., Hawes, A., Hendersen, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoque, M., Hollins, B., Howells, S., Huijk, S., Hume, J., Iddlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Liu, J., Kowis, C., Kraft, C. H., Lebow, H., Levan, J., Lewin, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshoewa, L., Loulsedeh, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaraine, M., Mahmoud, M., Mailly, K., Mangum, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Muniz, M., Myllyniemi, M., Naukkelius, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeleme, O., Okwuonu, G., Olarinpasagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, I., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puaco, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shartsbeyn, A., Sisson, I., Sitter, C. D., Smalls, D., Sned, A., Soderstrom, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Verfa, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.			
JOURNAL	Direct Submission			
REFERENCE	2 (bases 1 to 20953)			
2 (bases 1 to 20953)				

JOURNAL Submitted (15-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KCBA

Center clone name: CH230-99D23

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap: version 0.990329

Consensus quality: 101023 bases at least Q40

Consensus quality: 113373 bases at least Q30

Consensus quality: 122379 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length!
 *(see http://www.hgsc.bcm.tmc.edu/docs/Genbank/draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently consists of 86 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1037: contig of 1037 bp in length

* 1137: gap of unknown length

* 1515: contig of 1515 bp in length

* 2652: gap of unknown length

* 2752: gap of unknown length

* 2753: contig of 1561 bp in length

* 4313: contig of 1375 bp in length

* 4413: gap of unknown length

* 4414: contig of 1498 bp in length

* 5912: gap of unknown length

* 6011: contig of 1030 bp in length

* 7041: gap of unknown length

* 7142: 8516: contig of 1375 bp in length

* 8517: 8616: gap of unknown length

* 8617: 10009: contig of 1393 bp in length

* 10010: 10109: gap of unknown length

* 11247: contig of 1138 bp in length

* 11348: 12559: contig of 1212 bp in length

* 12569: gap of unknown length

* 12660: 14359: contig of 1700 bp in length

* 14360: 14459: gap of unknown length

* 15563: contig of 1104 bp in length

* 15564: 15663: gap of unknown length

* 15664: 17284: contig of 1621 bp in length

* 17285: 17384: gap of unknown length

* 17385: 18985: contig of 1601 bp in length

* 18986: 19085: gap of unknown length

* 20583: contig of 1498 bp in length

* 20584: 20683: gap of unknown length

* 20684: 22246: contig of 1563 bp in length

* 22247: 23478: contig of 1132 bp in length

* 23479: 23578: gap of unknown length

* 23579: 25234: contig of 1656 bp in length

* 25235: 25334: gap of unknown length

* 26448: contig of 1114 bp in length

* 26549: 28198: contig of 1650 bp in length

* 28199: 28298: gap of unknown length

* 28299: 29636: contig of 1338 bp in length

* 29637: 29736: gap of unknown length

* 31407: 31506: gap of unknown length

* 32893: contig of 1387 bp in length

* 32993: gap of unknown length

Query Match 73.9%; Score 17; DB 2; Length 202953;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCAATCTTACTTT 17

QD 191355 GGGCAATCTTACTTT 191339

RESULT 9
 AC124772
 LOCUS
 DEFINITION MUS musculus chromosome UNK clone RP23-47P3, WORKING DRAFT
 ACCESSION AC124772
 VERSION AC124772.1 GI:2142045
 KEYWORDS house mouse; HGSC; PHASE1; HGSC_DRAFT.
 SOURCE Mus musculus
 ORGANISM Bokaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 235301)
 AUTHORS McPherson, J. D. and Waterston, R. H.
 TITLE The sequence of *Mus musculus* clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 235301)
 AUTHORS McPherson, J. D. and Waterston, R. H.
 TITLE Direct Submission
 JOURNAL Submitted (15-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 COMMENT
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>
 Contact: submissions@wustl.edu
 Center project name: M_Ba0047203
 ----- Summary Statistics -----
 sequencing vector: M13; 0%
 Sequencing vector: plasmid; 100%
 Chemistry: dye-primer ET; 0% of reads
 Chemistry: dye-terminator Big dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 224743 bases at least Q40
 Consensus quality: 226669 bases at least Q30
 Consensus quality: 228045 bases at least Q20
 Insert size: 234906; sum-of-contigs
 Quality coverage: 8.20 in Q20 bases; agarose-1%
 Quality coverage: 6.79 in Q20 bases; sum-of-contigs
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 19 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1168: contig of 1168 bp in length
 * 1268: gap of unknown length
 * 2715: contig of 1447 bp in length
 * 2815: gap of unknown length
 * 4074: contig of 1259 bp in length
 * 4174: gap of unknown length
 * 4175: contig of 1777 bp in length
 * 5952: contig of 6051 bp in length
 * 6052: gap of unknown length
 * 8675: gap of unknown length
 * 8774: contig of 3910 bp in length
 * 12685: 12784: gap of unknown length
 * 12785: 15704: contig of 2920 bp in length
 * 15705: gap of unknown length
 ----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>
 Contact: submissions@wustl.edu
 Center project name: M_Ba0047203

 FEATURES
 source
 /organism="Mus musculus"
 /ab_xref="taxon:10090"
 /clone="RP23-47P3"
 misc_feature
 /note="assembly_name:Contig3"
 misc_feature
 /note="assembly_name:Contig15"
 misc_feature
 /note="assembly_name:Contig16"
 misc_feature
 /note="assembly_name:Contig17"
 misc_feature
 /note="assembly_name:Contig18"
 misc_feature
 /note="assembly_name:Contig19"
 misc_feature
 /note="assembly_name:Contig20"
 misc_feature
 /note="assembly_name:Contig21"
 misc_feature
 /note="assembly_name:Contig22"
 misc_feature
 /note="assembly_name:Contig23"
 misc_feature
 /note="assembly_name:Contig24"
 misc_feature
 /note="assembly_name:Contig25"
 misc_feature
 /note="assembly_name:Contig26"
 misc_feature
 /note="assembly_name:Contig27"
 misc_feature
 /note="assembly_name:Contig28"
 misc_feature
 /note="assembly_name:Contig29"
 misc_feature
 /note="assembly_name:Contig30"
 misc_feature
 /note="assembly_name:Contig31"
 misc_feature
 /note="assembly_name:Contig32"
 BASE COUNT 63617 a
 ORIGIN 52557 c 53697 g 63522 t 1928 others
 Query Match 73.9%; Score 17; DB 2; Length 235301;

QY	2	GGGCAATCTTACTTC 18		166595	239619: contig of 72925 bp in length.
Db	101622	GGGCAATCTTACTTC 101638		166695	239619: contig of 72925 bp in length.
RESULT	10				
ACCESSION	AC124601				
LOCUS	AC124601	239619 bp DNA linear HTGS 08-AUG-2002		112837	112936: gap of unknown length
DEFINITION	Mus musculus chromosome UNK clone RP23-27M3, WORKING DRAFT	sequence, 7 unordered pieces.	112937	112937	140005: contig of 2073 bp in length
VERSION	AC124601	AC124601.2 G1:22138700		112937	140005: contig of 2073 bp in length
KEYWORDS	HTGS; PHASE1; HTGS_DRAFT; HTGS_FULLTOP; house mouse; Mus musculus	htgs; htgs_phase1; htgs_draft; htgs_fulltop; house mouse; Mus musculus	112937	140005	gap of unknown length
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 239619)	organism="Mus musculus" /dbxref="taxon:00990" /chromosome="UNK" /clone="RP23-27M3"	112937	140005	contig of 26485 bp in length
REFERENCE	1 (bases 1 to 239619)	note="assembly_name:Contig23"	112937	140005	gap of unknown length
AUTHORS	McPherson, J.D. and Waterston, R.H.	note="assembly_name:Contig24"	112937	140005	contig of 72925 bp in length
TITLE	The sequence of <i>Mus musculus</i> clone Unpublished	note="assembly_name:Contig26"	112937	140005	contig of 72925 bp in length
JOURNAL		note="assembly_name:Contig27"	112937	140005	contig of 72925 bp in length
REFERENCE	2 (bases 1 to 239619)	note="assembly_name:Contig28"	112937	140005	contig of 72925 bp in length
AUTHORS	McPherson, J.D. and Waterston, R.H.	note="assembly_name:Contig28"	112937	140005	contig of 72925 bp in length
TITLE	Direct Submission	note="assembly_name:Contig28"	112937	140005	contig of 72925 bp in length
JOURNAL	Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA	note="assembly_name:Contig28"	112937	140005	contig of 72925 bp in length
COMMENT	3 (bases 1 to 239619)	note="assembly_name:Contig28"	112937	140005	contig of 72925 bp in length
Center: Washington University Genome Sequencing Center					
Center code: WUGSC					
Web site: http://genome.wustl.edu/gsc/index.shtml					
Contact: submissions@watson.wustl.edu					
----- Project Information -----					
Center project name: M_BA0027M03					
----- Summary Statistics -----					
Sequencing vector: M13; 0%					
Sequencing vector: plasmid; 100%					
Chemistry: dye-primer; 0% of reads					
Chemistry: dye-terminator; BiQ dye; 100% of reads					
Assembly program: Phrap; version 0.990319					
Consensus quality: 237244 bases at least 040					
Consensus quality: 237942 bases at least 030					
Consensus quality: 237926 bases at least 020					
Insert size: 213000; agarose-1p					
Insert size: 239019; sum-of-contigs					
Quality coverage: 11.66 in Q20 bases; aquause-fp					
Quality coverage: 9.65 in Q20 bases; sum-of-contigs					
----- This is 'working draft' sequence. It currently					
----- consists of 7 contigs. The true order of the pieces					
----- is not known and their order in this sequence record is					
----- arbitrary. gaps between the contigs are represented as					
----- runs of N, but the exact sizes of the gaps are unknown.					
----- This record will be updated with the finished sequence					
----- as soon as it is available and the accession number will					
----- be preserved.					
----- 1687: contig of 1687 bp in length					
----- 1688 1787: gap of unknown length					
----- 1788 7817: contig of 6030 bp in length					
----- 7818 7917: gap of unknown length					
----- 7918 9627: contig of 88310 bp in length					
----- 9628 96327: gap of unknown length					
----- 96328 112836: contig of 16509 bp in length					
FEATURES	source	misc_feature	1. .1687	112837	112837
BASE COUNT	274	a /dbxref="taxon:9606"	note="assembly_name:Contig28"	112837	112837
ORIGIN		note="assembly_name:Contig28"	note="assembly_name:Contig28"	112837	112837
Query Match		note="assembly_name:Contig28"	note="assembly_name:Contig28"	112837	112837
Best Local Similarity	69.6%	note="assembly_name:Contig28"	note="assembly_name:Contig28"	112837	112837
Matches	16; Conservative	note="assembly_name:Contig28"	note="assembly_name:Contig28"	112837	112837
FEATURES	source	misc_feature	1. .700	112837	112837
BASE COUNT	274	a /dbxref="taxon:9606"	note="assembly_name:Contig28"	112837	112837
ORIGIN		note="assembly_name:Contig28"	note="assembly_name:Contig28"	112837	112837
Query Match		note="assembly_name:Contig28"	note="assembly_name:Contig28"	112837	112837
Best Local Similarity	100.0%	note="assembly_name:Contig28"	note="assembly_name:Contig28"	112837	112837
Matches	16; Conservative	note="assembly_name:Contig28"	note="assembly_name:Contig28"	112837	112837
QY	2	GGGCAATCTTACTTC 17		140110	140109: gap of unknown length
Db	385	GGGCAATCTTACTTC 370		166694	gap of unknown length

RESULT 12				SOURCE	Homo sapiens.
LOCUS AF086808	2238 bp	mRNA	linear	MAM 06-OCT-1998	ORGANISM Homo sapiens.
DEFINITION Bos taurus fertilin beta (ADAM 2) mRNA, complete cds.					Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.
ACCESSION AF086808					
VERSION AF086808.1					
KEYWORDS					
ORGANISM	Bos taurus.				
REFERENCE 1	(bases 1 to 2238)				
AUTHORS Waters, S. I. and White, J. M.					
TITLE Biochemical and molecular characterization of bovine fertilin alpha and beta (ADAM 1 and ADAM 2): a candidate sperm-egg binding/fusion complex.					
JOURNAL Biol. Reprod.	56 (5), 1245-1254 (1997)				
MEDLINE 97304362					
PUBLMED 9160725					
REFERENCE 2	(bases 1 to 2238)				
AUTHORS Waters, S. I. and White, J. M.					
TITLE Direct Submission					
JOURNAL Submitted (25-AUG-1998) Department of Cell Biology, University of Virginia, Health Sciences Center, Box 439, Charlottesville, VA 22908, USA					
FEATURES Source					
gene	/product="Bos taurus"				
	/db_xref="taxon:9913"				
	/tissue_type="testis"				
	/protein_id="AA62753_1"				
	/db_xref="GI:3661512"				
CDS					
	1..2238				
	/gene="ADAM 2"				
	/codon_start=1				
	/product="fertilin beta"				
	/protein_id="AA62753_1"				
	/db_xref="GI:3661512"				
	/translation="MCILFLAIGLGLQTDNDSRLRVQPTPEKIRSTSGGGVETH				
	VSYIILIEGKTYVNLMQAKLPHNFRVYGGSGMKPPLHEFNCYQYQYLGYP				
	NSMAIISTCTGURGLQFENSYGPLEPLSIFGEMWYQKPKROSSSSYTERFEL				
	REPKKQVNPFLPDSQIEMHIVVKDLINHGMADTVIKQFLQFLGTTNAFTS				
	INIVTILSSILWIDENKTPYDANBLRHKWNRSSYLLRPHDMLFLWYRKS				
	YIGAFQGRMDKHYGGVHALHSITLESAVIAQLSISMGIPYDDINCKHCPGD				
	VCINPAPAVHSSKLFESVDEDFKIFSPKSDQLQNRDLPDYPKSAVCGNGAVE				
	EGEOQDCGNGKACDAPDTCVADTCRQFQPSACDGICGCCSCAPTPKGHCRGSTD				
	CDLHEYNGGSAACOBYVYDQHPCGONQCLTSGIVQDQTKOEDFEGCTSYPA				
	ECFORLNSHNDISGQCTPCTSENPKLITDQKREVSVENASVMSNQ				
	NGKTCIGLHLHBYGNFEDGMMWKGDKVCGESCTQNKQVNSFINDCNHCKNQ				
	VCNKHKHCHENPSYLPPNEISAPGWFQGSIISGNGNPPSPKPTGPAFTGTTPLAE				
	SRYENVYRSKPRTRMFFLEPFFILCLVLIATLVKVKYFQRKKWRTEDYISDEQLESE				
	SEPKD*				
BASE COUNT	682	a	397	c	495
ORIGIN					9
					664
					t
RESULT 13					
LOCUS BC030525/c	2558 bp	mRNA	linear	PRI 20-MAY-2002	
DEFINITION Homo sapiens, Similar to LOC201361, clone MGC:40524 IMAGE:5207811,					
ACCESSION BC030525					
VERSION BC030525.1					
KEYWORDS MGC.					
RESULT 14					
LOCUS AC004649	24102 bp	DNA	linear	PRI 01-MAY-1998	
DEFINITION Homo sapiens chromosome 5, P1 clone 603E12 (LBNL H26), complete sequence.					
ACCESSION AC004649	L77045	L77046	181623	L81782	L81783
VERSION AC004649.1					
KEYWORDS MGC.					

;/
;/ ADDRESSEE: Slim & McBurney
;/ STREET: Suite 701, 330 University Avenue
;/ CITY: Toronto
;/ STATE: Ontario
;/ COUNTRY: Canada
;/ ZIP: M5G 1R7
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: PatentIn Release #1.0, Version #1.30
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/549,515
;/ FILING DATE: 27-OCT-1995
;/ CLASSIFICATION: 536
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Stewart, Michael I
;/ REGISTRATION NUMBER: 24,973
;/ REFERENCE/DOCKET NUMBER: 1038-522
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (416) 595-1155
;/ TELEFAX: (416) 595-1163
;/ INFORMATION FOR SEQ ID NO: 1:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 4051 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ FEATURE:
;/ NAME/KEY: CDS
;/ LOCATION: join(65..2482, 2496..3110, 3115..3951)
;/ US-08-549-515-1

Query Match 56.5%; Score 13; DB 3; Length 4051;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 5 CAAATCTTACRTT 17
Db 138 CAAATCTTACRTT 150

Search completed: January 11, 2003, 23:01:58
Job time : 70 secs

Db 516 AAATCTTACCTTTC 504 US-08-549-515-2

RESULT 12 Query Match 56.5%; Score 13; DB 3; Length 2418;
US-09-134-001C-1731/c Best Local Similarity 100.0%; Pred. No. 14;
; Sequence 1731, Application US/09134001C Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; Patent No. 6380370

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO: 1731
; LENGTH: 2073
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis

Query Match 56.5%; Score 13; DB 4; Length 2073;
Best Local Similarity 100.0%; Pred. No. 13; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 CAAATCTTACCTT 17 Db 1595 CAAATCTTACCTT 1583

RESULT 13 Query Match 56.5%; Score 13; DB 3; Length 2418;
US-08-549-515-2 Best Local Similarity 100.0%; Pred. No. 14;
; Sequence 2, Application US/08549515 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; Patent No. 6054123

; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; TITLE OF INVENTION: Haemophilus Influenzae
; TITLE OF INVENTION: Dimethylsulphoxide Reductase Enzyme
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/549,515
FILING DATE: 27-OCT-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
TELECOMMUNICATION INFORMATION:
TELEPHONE: (116) 595-1155
TELEFAX: (116) 595-1163
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2418 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

Qy 5 CAAATCTTACCTT 17 Db 279 CAAATCTTACCTT 267

RESULT 14 Query Match 56.5%; Score 13; DB 4; Length 3083;
US-08-961-527-138/C Best Local Similarity 100.0%; Pred. No. 14;
; Sequence 138, Application US/08961527
; Patent No. 64,20135

; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/13
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 4,24
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brooks, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-5512
INFORMATION FOR SEQ ID NO: 138:
SEQUENCE CHARACTERISTICS:
LENGTH: 3083 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

Qy 5 CAAATCTTACCTT 17 Db 279 CAAATCTTACCTT 267

RESULT 15 Query Match 56.5%; Score 13; DB 4; Length 3083;
US-08-961-527-138/C Best Local Similarity 100.0%; Pred. No. 14;
; Sequence 1, Application US/08549515
; Patent No. 654123

; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; TITLE OF INVENTION: Haemophilus Influenzae
; TITLE OF INVENTION: Dimethylsulphoxide Reductase Enzyme
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:

ATTORNEY/AGENT INFORMATION:

NAME: Gimml, Edward R

REGISTRATION NUMBER: 38 891

REFERENCE/DOCKET NUMBER: P50475

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478

TELEX: 610-270-5090

TELEFAX: 610-270-4478

INFORMATION FOR SEQ ID NO: 167:

SEQUENCE CHARACTERISTICS:

LENGTH: 677 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLogy: linear

US-08-858-207A-167

Query Match 56.5%; Score 13; DB 4; Length 677;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

Query Match 56.5%; Score 13; DB 4; Length 765;

Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative

Organism: *Staphylococcus epidermidis*

US-09-134-001C-1728

TYPE: Nucleic Acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: polynucleotide
 US-08-365-981-1

Query Match Similarity 56.5%; score 13; DB 1; length 333;
 Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 AACATCTACTTG 19
 ||||||| 156 AACATCTACTTG 144

RESULT 5 US-09-134-001C-96
 Sequence 96, Application US/09134001C
 Patent No. 6380370

GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC-007
 CURRENT APPLICATION NUMBER: US/09-134, 001C
 CURRENT FILING DATE: 1998-08-13
 PRIORITY NUMBER: US 60/064, 964
 PRIOR FILING DATE: 1997-11-08
 PRIOR APPLICATION NUMBER: US 60/055, 779
 PRIORITY FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO 96
 LENGTH: 339
 TYPE: DNA
 ORGANISM: ~~Staphylococcus epidermidis~~

US-09-134-001C-96

Query Match Similarity 56.5%; Score 13; DB 4; Length 339;
 Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 CAAATCTACTT 17
 ||||||| 85 CAAATCTACTT 97

RESULT 6 US-09-134-001C-131
 Sequence 131, Application US/09134001C
 Patent No. 6380370

GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC-007
 CURRENT APPLICATION NUMBER: US/09-134, 001C
 CURRENT FILING DATE: 1998-08-13
 PRIORITY NUMBER: US 60/064, 964
 PRIOR FILING DATE: 1997-11-08
 PRIORITY FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO 131
 LENGTH: 339
 TYPE: DNA
 ORGANISM: ~~Staphylococcus epidermidis~~

US-09-134-001C-131

Db 85 CAAATCTACTT 97

RESULT 7 US-09-385-982-148/C
 Sequence 148, Application US/09385982
 Patent No. 6262334

GENERAL INFORMATION:
 APPLICANT: ENDEGE, WILSON O., ET AL.
 TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 TITLE OF INVENTION: PRODUCTS: 11
 FILE REFERENCE: CDDNA-260XX
 CURRENT APPLICATION NUMBER: US/09/385, 982
 CURRENT FILING DATE: 1999-08-30
 EARLIER APPLICATION NUMBER: 09/128, 111
 EARLIER FILING DATE: 1999-06-08
 EARLIER APPLICATION NUMBER: 60/117, 393
 EARLIER FILING DATE: 1999-01-27
 EARLIER APPLICATION NUMBER: 60/098, 639
 EARLIER FILING DATE: 1998-08-31
 NUMBER OF SEQ ID NOS: 544
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO 148
 LENGTH: 614
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: (1) .. (614)
 OTHER INFORMATION: n = A,T,C or G

US-09-385-982-148/C

Query Match Similarity 56.5%; Score 13; DB 4; Length 614;
 Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 CAAATCTACTT 17
 ||||||| 304 CAAATCTACTT 292

RESULT 8 US-08-858-207A-167/C
 Sequence 167, Application US/08858207A
 Patent No. 6348328

GENERAL INFORMATION:
 APPLICANT: HICK, Michael
 APPLICANT: Hodson, John
 APPLICANT: Knowles, David
 APPLICANT: Nicholas, Richard
 APPLICANT: Stodola, Robert
 TITLE OF INVENTION: NO. 6348328 e1 Compounds
 NUMBER OF SEQUENCES: 552
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Smithkline Beecham Corporation
 STREET: 709 Sweden Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406-0939
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/858, 207A
 FILING DATE: 09-MAY-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/017670
 FILING DATE: 14-MAY-1996

Qy 5 CAAATCTACTT 17

|||||||||
DB 21156 CTACTTCGCCTC 21169
RESULT 2
US-07-195-859B-10/c
; Sequence 10, Application US/07795859B
; Patent No. 542262
; GENERAL INFORMATION:
; APPLICANT: Anderson, Steffan
; TITLE OF INVENTION: Steroid 5'-Reductases
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White and Durkee
; STREET: P.O Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Floppy disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/795,859B
; FILING DATE: 18-NOV-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:260/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 320-7200
; TELEFAX: (512) 474-7677
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; TELEPHONE: (512) 474-7677
; US-08-457-616-10
; Query Match 56.5%; score 13; DB 1; length 100;
; Best Local Similarity 100.0%; Pred. No. 11; 0; Mismatches 0; Indels 0; Gaps 0;
; Matches 13; Conservative 0;
; OY 6 AAATCTTACTTC 18
; |||||||
; Db 16 AAATCTTACTTC 4
; |||||||
RESULT 4
US-08-365-981-1/c
; Sequence 1, Application US/08365981
; Patent No. 5583030
; GENERAL INFORMATION:
; APPLICANT: Robert DICKSON et al
; TITLE OF INVENTION: METHOD FOR OBTAINING
; TITLE OF INVENTION: ANTI-FUNGAL AND HERBICIDAL COMPOUNDS THAT TARGET THE
; TITLE OF INVENTION: FIRST COMMITTED STEP IN SPHINGOLIPID LONG-CHAIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LOWE, PRICE, LFRLANC & BECKER
; STREET: SITE 300, 99 CANAL CENTER PLAZA
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: DOS Text File
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,981
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/906,899
; FILING DATE: 06/30/92
; ATTORNEY/AGENT INFORMATION:
; NAME: D.J. MILLS
; REGISTRATION NUMBER: 34506
; REFERENCE/DOCKET NUMBER: 434-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 684-1111
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333

GenCore version 5.1.3
 Copyright (c) 1993 - 2003 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: January 11, 2003, 22:56:01 ; Search time 58 Seconds
 (without alignments)
 121.613 Million cell updates/sec

Title: SEQ01-T-AT-12347_COPY_12336_12358
 Perfect score: 23
 Sequence: 1 99ggaatcttactttcgcc 23

Scoring table: OLIGO-NUC
 Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 12

Total number of hits satisfying chosen parameters: 122

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*

1: /rcgn2_6/ptodata/1/inna/5A_COMB.seq: *
 2: /rcgn2_6/ptodata/1/inna/5B_COMB.seq: *
 3: /rcgn2_6/ptodata/1/inna/6A_COMB.seq: *
 4: /rcgn2_6/ptodata/1/inna/6B_COMB.seq: *
 5: /rcgn2_6/ptodata/1/inna/9CTUS_COMB.seq: *
 6: /rcgn2_6/ptodata/1/inna/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT 1
 US-08-752-760A-1
 ; Sequence 1, Application US/08752760A
 ; Patent No. 5877011
 ; GENERAL INFORMATION:
 ; APPLICANT: Armentano, Donna
 ; APPLICANT: Gregory, Richard J.
 ; APPLICANT: Smith, Alan E.
 ; TITLE OF INVENTION: CHIMERIC ADENOVIRAL VECTORS
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Baker & Botts, I.L.P.
 ; STREET: 30 Rockefeller Plaza
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: U.S.A.
 ; ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08-752-760A
 FILING DATE: 20-NOV-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:
 NAME: Seide, Rochelle K
 REGISTRATION NUMBER: 32,300
 REFERENCE/DOCKET NUMBER: A31385
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-705-5000
 TELEFAX: 212-705-5020

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 35081 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-752-760A-1

Sequence 12, Appl
 Sequence 29, Appl
 Sequence 31, Appl
 Sequence 36, Appl
 Sequence 32, Appl
 Sequence 3, Appl
 Sequence 4, Appl
 Sequence 4, Appl
 Sequence 4, Appl
 Sequence 4, Appl
 Sequence 32, Appl
 Sequence 205, App
 Sequence 205, App
 Sequence 205, App
 Sequence 29, Appl
 Sequence 391, App
 Sequence 375, App
 Sequence 53, Appl

c 28 12 52.2 86 2 US-08-870-518-12
 c 29 12 52.2 255 2 US-08-870-518-29
 c 30 12 52.2 294 2 US-08-870-518-31
 c 31 12 52.2 296 4 US-09-370-838-86
 c 32 12 52.2 350 3 US-08-888-077A-32
 c 33 12 52.2 358 1 US-08-686-878A-3
 c 34 12 52.2 400 3 US-08-714-918-4
 c 35 12 52.2 400 4 US-09-265-315-4
 c 36 12 52.2 400 4 US-09-265-315-4
 c 37 12 52.2 400 4 US-09-266-417-4
 c 38 12 52.2 439 2 US-08-870-518-32
 c 39 12 52.2 505 4 US-08-991-789A-205
 c 40 12 52.2 505 4 US-09-062-451-205
 c 41 12 52.2 505 4 US-09-598-326-205
 c 42 12 52.2 515 4 US-09-586-719-29
 c 43 12 52.2 572 4 US-08-961-527-391
 c 44 12 52.2 611 4 US-09-328-111-375
 c 45 12 52.2 787 4 US-09-370-838-53

ALIGNMENTS

Sequence 32, Appl
 Sequence 205, App
 Sequence 205, App
 Sequence 205, App
 Sequence 29, Appl
 Sequence 391, App
 Sequence 375, App
 Sequence 53, Appl

Query Match 60.9%; Score 14; DB 2; Length 35081;
 Best Local Similarity 100.0%; Pred. No. 3.6; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0;

QY 10 CTTACTTCGCTCC 23

ALIGNMENT

FILE REFERENCE: 004374.00739 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 CURRENT APPLICATION NUMBER: US/10/149,256 ; FILE REFERENCE: SCRIP1300-3
 CURRENT FILING DATE: 2002-06-10 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 PRIORITY APPLICATION NUMBER: PCT/EP00/12110 ; CURRENT FILING DATE: 2001-08-24
 PRIORITY APPLICATION NUMBER: US 60/239,106 ; PRIORITY APPLICATION NUMBER: US 60/227,866
 PRIORITY FILING DATE: 2000-10-11 ; PRIORITY FILING DATE: 2001-01-16
 PRIORITY FILING DATE: 1999-12-08 ; PRIORITY APPLICATION NUMBER: US 60/300,111
 NUMBER OF SEQ ID NOS: 10 ; PRIORITY FILING DATE: 2001-06-22
 SOFTWARE: FastSEQ for Windows Version 4.0 ; NUMBER OF SEQ ID NOS: 5379
 SEQ ID NO 3 ; SEQ ID NO 395
 LENGTH: 630 ; LENGTH: 1695
 TYPE: DNA ; TYPE: DNA
 ORGANISM: Homo sapiens ; ORGANISM: Arabidopsis thaliana
 FEATURE: NAME/KEY: misc-feature
 LOCATION: (1)..(630) ; US-09-938-842A-395
 OTHER INFORMATION: n = A,T,C or G ; OTHER INFORMATION: n = A,T,C or G
 / US-10-149-256-3 ; US-09-938-842A-395
 /
 Query Match 65.2%; Score 15; DB 9; Length 630;
 Best Local Similarity 100.0%; Pred. No. 1.7; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 GGGCAATCTTACCT 16
 Db 338 GGGCAATCTTACCT 324
 /
 RESULT 3 ;
 US-09-974-300-3323 ;
 Sequence 3323, Application US/09974300 ;
 ; Patent No. US20020146721A1
 /
 GENERAL INFORMATION:
 /
 APPLICANT: Berka, Randy M.
 /
 APPLICANT: Clausen, Ib Groth
 /
 TITLE OF INVENTION: Methods For Monitoring Multiple Gene
 /
 TITLE OF INVENTION: Expression
 /
 FILE REFERENCE: 10085_500-05
 /
 CURRENT APPLICATION NUMBER: US/09/974,300
 /
 CURRENT FILING DATE: 2001-10-05
 /
 PRIORITY APPLICATION NUMBER: 09/630,598
 /
 PRIORITY FILING DATE: 2000-10-06
 /
 PRIORITY APPLICATION NUMBER: 60/279,526
 /
 PRIORITY FILING DATE: 2001-03-27
 /
 NUMBER OF SEQ ID NOS: 8481
 /
 SOFTWARE: FastSEQ for Windows Version 4.0
 /
 SEQ ID NO 3323 ;
 LENGTH: 660 ;
 /
 TYPE: DNA ;
 /
 ORGANISM: Bacillus licheniformis ;
 /
 US-09-974-300-3323 ;
 /
 Query Match 60.9%; Score 14; DB 10; Length 660;
 Best Local Similarity 100.0%; Pred. No. 7.1; Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 6 AACCTTACTTCC 19
 Db 371 AACCTTACTTCC 384
 /
 RESULT 4 ;
 US-09-938-842A-395 ;
 Sequence 395, Application US/09938842A ;
 ; Patent No. US2002016078A1
 /
 GENERAL INFORMATION:
 /
 APPLICANT: Harper, Jeff
 /
 APPLICANT: Krebs, Joel
 /
 APPLICANT: Wang, Xun
 /
 APPLICANT: Zhu, Tong
 /
 TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 /
 TITLE OF INVENTION: SAME, AND METHODS OF USE
 /
 FILE REFERENCE: SCRIP1300-3
 /
 CURRENT APPLICATION NUMBER: US/09/938,842A
 /
 CURRENT FILING DATE: 2001-08-24
 /
 PRIORITY APPLICATION NUMBER: US 60/227,866
 /
 PRIORITY FILING DATE: 2000-08-24
 /
 PRIORITY APPLICATION NUMBER: US 60/264,647
 /
 PRIORITY FILING DATE: 2001-01-16
 /
 PRIORITY APPLICATION NUMBER: US 60/300,111
 /
 PRIORITY FILING DATE: 2001-06-22
 /
 NUMBER OF SEQ ID NOS: 5379
 /
 SEQ ID NO 4537 ;
 LENGTH: 2000 ;
 /
 TYPE: DNA ;
 /
 ORGANISM: Arabidopsis thaliana ;
 /
 US-09-938-842A-4537 ;
 /
 Query Match 60.9%; Score 14; DB 9; Length 2000;
 Best Local Similarity 100.0%; Pred. No. 7.1; Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 5 CAACTCTTCTTC 18
 Db 1085 CAACTCTTCTTC 1098
 /
 RESULT 6 ;
 US-09-917-800A-1644/C ;
 Sequence 1644, Application US/09917800A ;
 ; Patent No. US20020119462A1
 /
 GENERAL INFORMATION:
 /
 APPLICANT: Mendick, Donna
 /
 APPLICANT: Porter, Mark
 /
 APPLICANT: Johnson, Kory
 /
 APPLICANT: Castle, Arthur
 /
 APPLICANT: Elashoff, Michael
 /
 APPLICANT: Gene Logic, Inc.

OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (3176)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (3186)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (3203)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (3217)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (3234)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (3412)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (3431)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (3478)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4342)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4351)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4360)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4361)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4375)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4435)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4445)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4459)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4470)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4472)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4475)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4486)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4503)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4507)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4515)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4524)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4558)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4616)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4705)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (4754)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (5311)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (5513)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (6363)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (7252)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (8375)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (8560)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (8878)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (9805)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (9839)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (10029)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (10046)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (10389)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 ;
 RESULT 9
 Query Match 60.9%; Score 14; DB 10; Length 17908;
 Best Local Similarity 100.0%; Pred. No. 73; Mismatches 0; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 CAATCTTACTTC 18
 Db 17586 CAATCTTACTTC 17573
 ;
 ; Sequence 26, Application US/09782378A
 ; Patent No. US20020102731A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hearing, Patrick
 ; APPLICANT: Bahou, Wadie
 ; APPLICANT: Sandal, Ziv

US-09-702-378A-26

APPLICANT: Gratenko, Dmitri
 TITLE OF INVENTION: Adenoviral Vectors
 FILE REFERENCE: 210121.497
 CURRENT APPLICATION NUMBER: US/09/867,701
 CURRENT FILING DATE: 2001-05-29
 NUMBER OF SEQ ID NOS: 10912
 PRIORITY FILING DATE: 60/237,747
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 9
 LENGTH: 248
 TYPE: DNA
 ORGANISM: Homo sapien
 US-09-867-701-9

TYPE: DNA
 ORGANISM: Human adenovirus type 17
 FEATURE: misc_feature
 NAME/KEY: misc_feature
 LOCATION: (25184)..(25184)
 OTHER INFORMATION: The n at this position can be a, c, t, or g.

US-09-782-378A 26

Query Match Best Local Similarity 60.9%; Score 14; DB 10; Length 35100; Matches 14; Conservative 100.0%; Pred. No. 7; 3; Indels 0; Gaps 0;

Qy 10 CTTACTTCTCGCTCC 23
 Db 21170 CTTACTTCTCGCTCC 21183

RESULT 10

Sequence 5826, Application US/09983965
 Patent No. US2002137160A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.
 APPLICANT: Tao, Ningbing
 APPLICANT: Bratt, John C.
 APPLICANT: Mathialagan, Nagappan
 TITLE OF INVENTION: NUCLEAR ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND FILE REFERENCE: 37-21(10297)C
 CURRENT APPLICATION NUMBER: US/09/983,965
 CURRENT FILING DATE: 2001-10-26

PRIOR APPLICATION NUMBER: US 09/465,231
 PRIOR FILING DATE: 1999-12-15
 PRIOR APPLICATION NUMBER: US 60/113,678
 NUMBER OF SEQ ID NOS: 5912
 SEQ ID NO 5826
 LENGTH: 197
 TYPE: DNA
 ORGANISM: Bos taurus
 FEATURE:
 OTHER INFORMATION: Clone ID: 60-LB34-059-Q1-F1-G12

US-09-983-965-5826

Query Match Best Local Similarity 100.0%; Score 13; DB 10; Length 197; Matches 13; Conservative 100.0%; Pred. No. 29; Indels 0; Gaps 0;

Qy 5 CAATCTCACTT 17
 Db 145 CAATCTCACTT 133

RESULT 11

Sequence 9, Application US/09867701
 Patent No. US20020132237A1

GENERAL INFORMATION:

APPLICANT: Aglate, Paul A.
 APPLICANT: Jones, Robert
 APPLICANT: Harlocker, Susan L.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

US-09-867-701-9

Query Match Best Local Similarity 56.5%; Score 13; DB 10; Length 264; Matches 13; Conservative 100.0%; Pred. No. 29; Indels 0; Gaps 0;

Qy 1 GGGCCAATCTTA 13
 Db 172 GGGCCAATCTTA 184

RESULT 13

Sequence 782, Application US/09922217
 Patent No. US2002007614A1

GENERAL INFORMATION:

RESULT 15
 US-09-867-701-4322/C
 ; Sequence 4322, Application US/09867701
 ; Patent No. US20020132237A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ajlate, Paul A.
 ; APPLICANT: Jones, Robert
 ; APPLICANT: Harlocke, Susan L.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; AND DIAGNOSTICS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.497
 ; CURRENT APPLICATION NUMBER: US/09/867,701
 ; CURRENT FILING DATE: 2001-05-29
 ; NUMBER OF SEQ ID NOS: 10912
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 4322
 ; LENGTH: 412
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE: misc_feature
 ; NAME/KEY: misc_feature
 ; LOCATION: 47, 69, 195, 262
 ; OTHER INFORMATION: n = A,T,C or G

Query Match 56.5%; Score 13; DB 10; Length 264;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGCAATCTTA 13
 Db 172 GGGGCAATCTTA 184

RESULT 14
 US-09-833-263-782
 ; Sequence 782, Application US/09833263
 ; Patent No. US2002010547A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Meagher, Madeleine J.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
 ; DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.471C12
 ; CURRENT APPLICATION NUMBER: US/09/833,263
 ; CURRENT FILING DATE: 2001-04-10
 ; NUMBER OF SEQ ID NOS: 1093
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 782
 ; LENGTH: 264
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE: misc_feature
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(264)
 ; OTHER INFORMATION: n = A,T,C or G

Query Match 56.5%; Score 13; DB 10; Length 264;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGCAATCTTA 13
 Db 172 GGGGCAATCTTA 184

Search completed: January 12, 2003, 00:28:41
 Job time : 73 secs